

GenCore version 5.1.5

Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 15:56:54 ; Search time 1090.13 Seconds
(without alignments)
11900.029 Million cell updates/sec

Title: US-09-434-382-28_COPY_21800_22600

Perfect score: 801

Sequence: 1 agtgcctgctctgtattttt.....agcgaagctttgaccggat 801

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum Match 100%

Listing first 45 summaries

Post-processing: Minimum Match 0%

Maximum DB seq length: 2000000000

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_nam:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rod:*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	354	44.2	356	9	AI475638 tc86d11.x
2	146.4	18.3	559	14	BM690352 UI-B-C10-
3	145.8	18.2	482	14	R55841 Yg89d01.r1
4	145.8	18.2	664	10	BE383336 601298249
5	145.8	18.2	676	10	BE260495 601150702
6	145.8	18.2	692	10	BE382353 601298656

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

7	145.8	18.2	698	12	BE795434
8	145.8	18.2	726	13	BM009099
9	145.8	18.2	756	13	BE795434
10	145.8	18.2	761	12	BE795820
11	145.8	18.2	874	14	BE795820
12	145.8	18.2	876	14	BE795820
13	145.8	18.2	884	12	BE795820
14	145.8	18.2	915	12	BE795820
15	145.8	18.2	915	14	BE795820
16	145.8	18.2	928	14	BE795820
17	145.8	18.2	960	12	BE795820
18	145.8	18.2	966	12	BE795820
19	145.8	18.2	980	14	BE795820
20	145.8	18.2	983	14	BE795820
21	145.8	18.2	1079	14	BE795820
22	145.8	18.2	1123	14	BE795820
23	145.8	18.2	1282	14	BE795820
24	145.4	18.2	898	14	BE795820
25	144.8	18.1	495	14	BE795820
26	144.4	18.0	614	12	BE795820
27	144.2	18.0	771	12	BE795820
28	144.2	18.0	791	12	BE795820
29	141.6	17.7	297	10	BE795820
30	140.2	17.5	992	12	BE795820
31	134.8	16.8	472	14	BE795820
32	134.8	16.8	1065	13	BE795820
33	133.8	16.7	735	12	BE795820
34	133.8	16.7	947	12	BE795820
35	133.8	16.7	1497	12	BE795820
36	133.2	16.6	341	14	BE795820
37	131.8	16.5	315	14	BE795820
38	126.6	15.8	930	14	BE795820
39	126.6	15.8	970	14	BE795820
40	124.6	15.6	745	12	BE795820
41	117.6	14.7	1002	14	BE795820
42	114.4	14.3	547	12	BE795820
43	114	14.2	269	12	BE795820
44	111.4	13.9	543	10	BE795820
45	108.6	13.6	713	10	BE795820

ALIGNMENTS

RESULT 1

AI475638

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AI475638 356 bp mRNA linear EST 18-MAR-1999
tc86d11.x1 NCI_CGAP_CLL1 Homo sapiens cDNA clone IMAGE:2073045 3',
mRNA sequence.

AI475638
AI475638.1 GI:4328683
EST.
human.

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 356)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov

M.D., Louis M. Staudt, M.D., Ph.D.
CNA Library Preparation: M. Bento Soares, Ph.D.
CNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 2499 Std Error: 0.00
Seq primer: -40UP from Gibco

High quality sequence stop: 351.
 Location/Qualifiers
 1..356
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2073045"
 /clone_lib="NCI CGAP CLL1"
 /tissue_type="B-cell, chronic lymphocytic leukemia"
 /lab_host="DH10B"
 /note="Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCGCATTTGTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized, and was constructed by Bento Soares and M. Fatima Bonaldo."
 BASE COUNT 83 a 85 c 86 g 102 t
 ORIGIN
 Query Match 44.2%; Score 354; DB 9; Length 356;
 Best Local Similarity 99.4%; Pred. No. 1.6e-62;
 Matches 354; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 115 TCATGCCAGCCTTTGTGAAGTCTACTTGAAGGGTTTCAGGGTCCATGATCTTCT 174
 DB 1 TGATGCCAGCCTTTGTGAAGTCTACTTGAAGGGTTTCAGGGTCCATGATCTTCT 60
 QY 175 TTGCTATAAGAGGATGACACATGTAATACCTTTATGTTAAATTAATGGCTTTTA 234
 DB 61 TTGCTATAAGAGGATGACACATGTAATACCTTTATGTTAAATTAATGGCTTTTA 120
 QY 235 TATTAGTCTCTCAAGCAAGCAGGAGACAGAGAATTTCTCAGTGTCTTCTTGCTCT 294
 DB 121 TATTAGTCTCTCAAGCAAGCAGGAGACAGAGAATTTCTCAGTGTCTTCTTGCTCT 180
 QY 295 GTCCAAAGCAGATCAGCTCTGACCATCAGAGTCTTCTAGTGGCAGTCTCTCT 354
 DB 181 GTCCAAAGCAGATCAGCTCTGACCATCAGAGTCTTCTAGTGGCAGTCTCTCT 240
 QY 355 TCCTCTCTCTTCTGAGCCCGACACGCTCTGCTACTGGACTGTGGTGGGGCACRTT 414
 DB 241 TCCTCTCTCTTCTGAGCCCGACACGCTCTGCTACTGGACTGTGGTGGGGCACATT 300
 QY 415 TGGCAGCTGTGCCGTCAATTACGAGACAGGTGGACAGGGTCTTGGGCACCCCTGG 470
 DB 301 TGGCAGCTGTGCCGTCAATTACGAGACAGGTGGACAGGGTCTTGGGCACCCCTGG 356
 RESULT 2
 BM690352 559 bp mRNA linear EST 28-FEB-2002
 LOCUS UI-E-C10-aah-b-03-0-UI.r1 UI-E-C10 Homo sapiens cDNA clone
 DEFINITION UI-E-C10-aah-b-03-0-UI 5', mRNA sequence.
 ACCESSION BM690352
 VERSION BM690352.1 GI:19003610
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 559)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 9704477
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250

Fax: 319 335 9565
 Email: meares@blue.weeg.uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
 Seq primer: M13 Reverse.
 Location/Qualifiers
 1..559
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="UI-E-C10-aah-b-03-0-UI"
 /clone_lib="UI-E-C10"
 /tissue_type="RPE and Choroid"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-C10 is a cDNA library containing the following tissue(s): RPE and Choroid. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."
 BASE COUNT 113 a 154 c 169 g 123 t
 ORIGIN
 Query Match 18.3%; Score 146.4; DB 14; Length 559;
 Best Local Similarity 89.7%; Pred. No. 4.1e-20;
 Matches 156; Conservative 1; Mismatches 17; Indels 0; Gaps 0;
 QY 339 GTGGCAGTGACTCTTCTCTCTTCTTCTGACCCCGACACGCTCTCTGTCTGGACT 398
 DB 7 GTCAAAATGTCAAGTCCACACTTGTCAACATAAGCCCGACACGCTCTCTGTCTGGACT 66
 QY 399 GTGTGAGGGACATTTGGGAGCTGTCCGTCATTACGGAGACAGGTGGACAGGTGCC 458
 DB 67 GTGTGAGGGACATTTGGGAGCTGTCCGTCATTACGGAGACAGGTGGACAGGTGCC 126
 QY 459 TGGGCACCTCGCTGTGTTGTGTCACCTGCACGACATCACCACCG 512
 DB 127 TGGGCACCTCGCTGTGTTGTGTCACCTGCACGACATCACCACCG 180
 RESULT 3
 R55841 482 bp mRNA linear EST 23-MAY-1995
 LOCUS Y989d01.r1 Soares infant brain INIB Homo sapiens cDNA clone
 DEFINITION IMAGE:40931 5', similar to SP:YK99_YEAST P36159 HYPOPHYSICAL 96.8 KD PROTEIN IN SIS2-MTD1 INTERGENIC ;, mRNA sequence.
 ACCESSION R55841
 VERSION R55841.1 GI:825947
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 482)
 AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
 TITLE The WashU-Merck EST Project
 JOURNAL Unpublished (1995)

BASE COUNT

[illegible]

BASE COUNT 199 a 251 c 249 g 185 t

ORIGIN

NOTE: THIS IS A NIN-MEC LIBRARY.

Qy 398

Qy 398 TGTGGTGAGGGCACTTTGGGCAGCTGTGCCGTCAATTACGGAGACCAGGTGGACAGGGTC 457
|||||:|||||
Db 473 TGTGGTGAGGGCACTTTGGGCAGCTGTGCCGTCAATTACGGAGACCAGGTGGACAGGGTC 532

Search completed: May 17, 2003, 18:18:28
Job time : 1096.13 secs

GenCore version 5.1.1.5

Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 17:27:39 ; Search time 101.551 Seconds
(without alignments)
10174.595 Million cell updates/sec

Title: US-09-434-382-28_COPY_21800_22600

Perfect score: 801

Sequence: 1 agtgcctctctgtattttt.....agcggaagctttgaccggat 801

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 810007 seqs, 644969091 residues

Total number of hits satisfying chosen parameters: 1620014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	800.6	100.0	26664	9	US-09-988-626-28
2	800.6	100.0	26664	9	US-09-988-687-28
3	145.8	18.2	2481	9	US-09-988-626-1
4	145.8	18.2	2481	9	US-09-988-687-1
5	145.8	18.2	2892	9	US-09-988-626-225
6	145.8	18.2	2892	9	US-09-988-687-225
7	145.8	18.2	2908	9	US-09-988-626-223
8	145.8	18.2	2908	9	US-09-988-687-223
9	145.8	18.2	2958	9	US-09-988-626-3
10	145.8	18.2	2958	9	US-09-988-687-3
11	138.6	17.3	139	9	US-09-988-626-20
12	138.6	17.3	139	9	US-09-988-687-20
13	102.2	12.8	2470	9	US-09-988-626-221
14	102.2	12.8	2470	9	US-09-988-687-221
15	39.2	4.9	1152	9	US-10-092-154-1517
16	39.2	4.9	1152	10	US-09-764-847-1517
17	39.2	4.9	1555	9	US-10-092-154-1518
18	39.2	4.9	1555	10	US-09-764-847-1518
19	38.8	4.8	1732	9	US-10-091-438-28

20	38.8	4.8	1732	10	US-09-764-853-379
21	38	4.7	10286	9	US-10-239-676-13
22	36.8	4.6	671	9	US-10-092-154-236
23	36.8	4.6	671	10	US-09-764-847-296
24	36.8	4.6	1578	10	US-09-822-830A-583
25	36	4.5	1182	10	US-09-922-261-261
c 25	36	4.5	2454	9	US-10-050-704-93
c 26	36	4.5	2664	9	US-09-809-391-255
c 27	36	4.5	512	9	US-09-918-995-24438
c 28	35.8	4.5	2005	10	US-09-887-576-50
29	35.8	4.5	8801	9	US-10-239-676-144
30	35.8	4.5	17056	10	US-09-833-238-1
c 31	35.8	4.5	2000	9	US-09-938-842A-3132
c 32	35.6	4.4	3253	10	US-09-759-359A-1
33	35.2	4.4	499	9	US-10-092-154-1546
34	35	4.4	499	10	US-09-764-847-1546
35	35	4.4	416	9	US-09-918-995-3035
c 36	34.8	4.3	416	9	US-09-918-995-16146
c 37	34.8	4.3	559	9	US-09-918-995-16146
c 38	34.4	4.3	478	9	US-09-918-995-16997
c 39	34.4	4.3	1119	9	US-10-123-155-352
c 40	34.4	4.3	513509	9	US-09-754-853A-4
c 41	34.2	4.3	419	10	US-09-960-352-13938
c 42	34	4.2	653	9	US-10-184-644-402
c 43	34	4.2	653	9	US-10-184-644-402
c 44	34	4.2	802	9	US-10-184-644-312
45	34	4.2	802	9	US-10-184-634-312

ALIGNMENTS

RESULT 1

US-09-988-626-28
; Sequence 28, Application US/09988626
; Publication No. US2003004959A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,626
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 26664
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (910)..(13104)
; OTHER INFORMATION: exon 1: 910-1154; exon 2: 1736-1786; exon 3:
; OTHER INFORMATION: 1925-1995; exon 4: 3025-3089; exon 5: 4361-4418;
; OTHER INFORMATION: exon 6: 5582-5650; exon 7: 7075-7194; exon 8:
; OTHER INFORMATION: 8186-8244; exon 9: 12878-12936; exon 10:
; OTHER INFORMATION: 13032-13104;
; NAME/KEY: misc_feature
; LOCATION: (13756)..(22917)
; OTHER INFORMATION: exon 11: 13756-13868; exon 12: 15283-15378; exon
; OTHER INFORMATION: 13: 16278-16416; exon 14: 16498-16583; exon 15:
; OTHER INFORMATION: 15883-18701; exon 16: 20349-20445; exon 17:
; OTHER INFORMATION: 22172-22310; exon 18: 22879-22917
; NAME/KEY: misc_feature

```
; LOCATION: (23045)..(26452)
; OTHER INFORMATION: exon 19: 23045-23154; exon 20: 23795-23895; exon
; OTHER INFORMATION: 21: 23973-24093; exon 22: 24354-24432; exon 23:
; OTHER INFORMATION: 25026-25170; exon 24: 25812-26036; polyadenylation
; OTHER INFORMATION: signal: 26447-26452
; NAME/KEY: variation
; LOCATION: (826)..(23879)
; OTHER INFORMATION: s at positions 826 and 23180 is G or C; y at
; OTHER INFORMATION: positions 1914, 5568, 7165, 16431, 1857 and 20486
; OTHER INFORMATION: is C or T; n at position 13128 is t or tgat; r at
; OTHER INFORMATION: positions 22211 and 23879 is A or G.
US-09-988-626-28

Query Match      100.0%; Score 800.6; DB 9; Length 26664;
Best Local Similarity 100.0%; Pred. No. 6.4e-234;
Matches 801; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGTGCTGTCTGGTATTTTCAACAGAGGCTGTGCCACAGTCAATCTGCATGGTCAGAT 60
Db 1800 AGTGCTGTCTGGTATTTTCAACAGAGGCTGTGCCACAGTCAATCTGCATGGTCAGAT 21859

Qy 61 TCATTGTTAGGACTAAATGCTTTAAGCCTCCTATAAACTTTTTTTTTTTTTTTGATGC 120
Db 21860 TCATTGTTAGGACTAAATGCTTTAAGCCTCCTATAAACTTTTTTTTTTTTTTTGATGC 21919

Qy 121 CCAGCCTTTGTGTAAGTCTACTTGAAGGGTTTCAGGGTTCATGGATACCTCTTTGCTA 180
Db 21920 CCAGCCTTTGTGTAAGTCTACTTGAAGGGTTTCAGGGTTCATGGATACCTCTTTGCTA 21979

Qy 181 TAAAGAGGATGACATGTAATCACTTTATGTTTAAATTAATTTGGCTTTTATATTAG 240
Db 21980 TAAAGAGGATGACATGTAATCACTTTATGTTTAAATTAATTTGGCTTTTATATTAG 22039

Qy 241 CTCCTCAAAGCAAAGCAGGAGACAGAAATTTTCAGATGCTCTTCTGGTCTCTGCTCA 300
Db 22040 CTCCTCAAAGCAAAGCAGGAGACAGAAATTTTCAGATGCTCTTCTGGTCTCTGCTCA 22099

Qy 301 AGCAGACATCAGCTCTGACCATCAGCATCTCTCTAGTGCAGTCACTCTCTCTCT 360
Db 22100 AGCAGACATCAGCTCTGACCATCAGCATCTCTCTAGTGCAGTCACTCTCTCTCT 22159

Qy 361 TCTCTTCTGACGCCGACACAGTCTCTGCTACTGACTGTGTGAGGGACATTTTGGGCA 420
Db 22160 TCTCTTCTGACGCCGACACAGTCTCTGCTACTGACTGTGTGAGGGACATTTTGGGCA 22219

Qy 421 GCTGTGCCGTCAATTACGGAGACAGGTGACAGGGTCTCTGGCACCCCTGGCTGTGT 480
Db 22220 GCTGTGCCGTCAATTACGGAGACAGGTGACAGGGTCTCTGGCACCCCTGGCTGTGT 22279

Qy 481 TGTGTCCACCTGCACGAGATCACCACAGGTGAGTGTGGGCTGGACCAAAAGCTGG 540
Db 22280 TGTGTCCACCTGCACGAGATCACCACAGGTGAGTGTGGGCTGGACCAAAAGCTGG 22339

Qy 541 ASCCTGGAGGAGGACTGCGACGTTGAGTTGGCCCTTTGGCTGCGTCTTTTCTCCGCT 600
Db 22340 ASCCTGGAGGAGGACTGCGACGTTGAGTTGGCCCTTTGGCTGCGTCTTTTCTCCGCT 22399

Qy 601 CCAAACTTGCCCGCAGCTTTTGTACTCTCTGCTAGGAAATGGTTTTTGGCAAAAC 660
Db 22400 CCAAACTTGCCCGCAGCTTTTGTACTCTCTGCTAGGAAATGGTTTTTGGCAAAAC 22459

Qy 661 TCAACATAGTCTCTGCGCCACAGAAATGTCTTCTCTCTGTTTCAGTCTCTTCTCTGC 720
Db 22460 TCAACATAGTCTCTGCGCCACAGAAATGTCTTCTCTCTGTTTCAGTCTCTTCTCTGC 22519

Qy 721 AGCAGACAGGTTTGTAGTTTACCGACCTTCTCTGAGTCTTCAATCTCACAGGCTGT 780
Db 22520 AGCAGACAGGTTTGTAGTTTACCGACCTTCTCTGAGTCTTCAATCTCACAGGCTGT 22579

Qy 781 CAGCGAAGCTTTGACCGGAT 801
Db 22580 CAGCGAAGCTTTGACCGGAT 22600
```

RESULT 2

```
US-09-988-687-28
; Sequence 28, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
; APPLICANT: Tavtighian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,687
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 26664
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (910)..(13104)
; OTHER INFORMATION: exon 1: 910-1154; exon 2: 1736-1786; exon 3:
; OTHER INFORMATION: 1925-1995; exon 4: 3025-3089; exon 5: 4361-4418;
; OTHER INFORMATION: exon 6: 5582-5650; exon 7: 7075-7194; exon 8:
; OTHER INFORMATION: 8186-8244; exon 9: 12878-12936; exon 10:
; OTHER INFORMATION: 13032-13104;
; NAME/KEY: misc feature
; LOCATION: (13756)..(22917)
; OTHER INFORMATION: exon 11: 13756-13868; exon 12: 15283-15378; exon
; OTHER INFORMATION: 13: 16278-16416; exon 14: 16498-16583; exon 15:
; OTHER INFORMATION: 18583-18701; exon 16: 20349-20445; exon 17:
; OTHER INFORMATION: 22172-22310; exon 18: 22879-22917
; NAME/KEY: misc feature
; LOCATION: (23045)..(26452)
; OTHER INFORMATION: exon 19: 23045-23154; exon 20: 23795-23895; exon
; OTHER INFORMATION: 21: 23973-24093; exon 22: 24354-24432; exon 23:
; OTHER INFORMATION: 25026-25170; exon 24: 25812-26036; polyadenylation
; OTHER INFORMATION: signal: 26447-26452
; NAME/KEY: variation
; LOCATION: (826)..(23879)
; OTHER INFORMATION: s at positions 826 and 23180 is G or C; y at
; OTHER INFORMATION: positions 1914, 5568, 7165, 16431, 1857 and 20486
; OTHER INFORMATION: is C or T; n at position 13128 is t or tgat; r at
; OTHER INFORMATION: positions 22211 and 23879 is A or G.
US-09-988-687-28
```

Query Match 100.0%; Score 800.6; DB 9; Length 26664;

Best Local Similarity 100.0%; Pred. No. 6.4e-234;

Matches 801; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 AGTGCTGTCTGGTATTTTCAACAGAGGCTGTGCCACAGTCAATCTGCATGGTCAGAT 60
Db 21800 AGTGCTGTCTGGTATTTTCAACAGAGGCTGTGCCACAGTCAATCTGCATGGTCAGAT 21859

Qy 61 TCATTGTTAGGACTAAATGCTTTAAGCCTCCTATAAACTTTTTTTTTTTTTTTGATGC 120
Db 21860 TCATTGTTAGGACTAAATGCTTTAAGCCTCCTATAAACTTTTTTTTTTTTTTTGATGC 21919

Qy 121 CCAGCCTTTGTGTAAGTCTACTTGAAGGGTTTCAGGGTTCATGGATACCTCTTTGCTA 180
Db 21920 CCAGCCTTTGTGTAAGTCTACTTGAAGGGTTTCAGGGTTCATGGATACCTCTTTGCTA 21979

Qy 181 TAAAGAGGATGACATGTAATCACTTTATGTTTAAATTAATTTGGCTTTTATATTAG 240
```



```

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51)..(2531)
; OTHER INFORMATION: coding sequence as in SEQ ID NO:1
US-09-988-626-3

Query Match      18.2%; Score 145.8; DB 9; Length 2958;
Best Local Similarity 89.1%; Pred. No. 7.6e-34;
Matches 156; Conservative 1; Mismatches 18; Indels 0; Gaps 0

QY 338 AGTGGCAGTGA CTCTTCCTCTCTCTCTCTGCGAGCCCGACACGCTCTCTGCTACTGGAC 397
Db 1536 ATTGGAATGTCAGTGCACACACTTGTCAACATAGACCCCGACACGCTCTCTGCTACTGGAC 1595

QY 398 TGTGGTAGGGGCACR TTTGGGCAGCTGTGCCGTCAATTACGGAGACCGAGTGACACAGGGTC 457
Db 1596 TGTGGTAGGGGCACAT TTTGGGCAGCTGTGCCGTCAATTACGGAGACCGAGTGACACAGGGTC 1655

QY 458 CTGGGCAACCCCTGGCTGTGTTGTGTCTCCACCTGCGACGAGATCACACACGG 512
Db 1656 CTGGGCAACCCCTGGCTGTGTTGTGTCTCCACCTGCGACGAGATCACACACGG 1710

RESULT 10
US-09-988-687-3
; Sequence 3, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
; APPLICANT: Tavtighian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,687
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2958
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51)..(2531)
; OTHER INFORMATION: coding sequence as in SEQ ID NO:1
US-09-988-687-3

Query Match      18.2%; Score 145.8; DB 9; Length 2958;
Best Local Similarity 89.1%; Pred. No. 7.6e-34;
Matches 156; Conservative 1; Mismatches 18; Indels 0; Gaps 0

QY 338 AGTGGCAGTGA CTCTTCCTCTCTCTCTGCGAGCCCGACACGCTCTCTGCTACTGGAC 397
Db 1536 ATTGGAATGTCAGTGCACACACTTGTCAACATAGACCCCGACACGCTCTCTGCTACTGGAC 1595

QY 398 TGTGGTAGGGGCACR TTTGGGCAGCTGTGCCGTCAATTACGGAGACCGAGTGACACAGGGTC 457
Db 1596 TGTGGTAGGGGCACAT TTTGGGCAGCTGTGCCGTCAATTACGGAGACCGAGTGACACAGGGTC 1655

QY 458 CTGGGCAACCCCTGGCTGTGTTGTGTCTCCACCTGCGACGAGATCACACACGG 512
Db 1656 CTGGGCAACCCCTGGCTGTGTTGTGTCTCCACCTGCGACGAGATCACACACGG 1710

```


GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 15:56:54 ; Search time 45.0714 Seconds
(without alignments)
5450.193 Million cell updates/sec

Title: US-09-434-382-28_COPY_21800_22600

Perfect score: 801

Sequence: 1 agtgcctgtcctgtatttt.....agcggagctttgacccggt 801

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	800.6	100.0	26664	4	US-09-564-805-28
2	145.8	18.2	2481	4	US-09-564-805-1
3	145.8	18.2	2892	4	US-09-564-805-225
4	145.8	18.2	2908	4	US-09-564-805-223
5	145.8	18.2	2958	4	US-09-564-805-3
6	138.6	17.3	139	4	US-09-564-805-20
7	102.2	12.8	2470	4	US-09-564-805-221
8	39.8	5.0	7218	1	US-08-232-463-14
9	39.4	4.9	840	4	US-09-376-728-1
10	36.8	4.6	810	4	US-09-134-001C-1951
11	36.6	4.6	19124	2	US-08-487-826B-13
12	36.6	4.6	4403765	4	US-09-103-840A-2
13	36.6	4.6	4411529	4	US-09-103-840A-1
14	36	4.5	1182	4	US-09-461-697-261
15	36	4.5	2664	4	US-09-149-476-255
16	36	4.5	50000	4	US-09-146-053-4
17	35.8	4.5	17056	4	US-09-245-041-3
18	35	4.4	289	4	US-09-007-005-17
19	35	4.4	289	4	US-09-244-796-17
20	34	4.2	3785	4	US-09-888-718-1
21	34	4.2	5852	4	US-09-853-768-10
22	34	4.2	7037	4	US-09-853-768-3
23	32.8	4.1	6999	1	US-08-276-594A-1
24	32.8	4.1	7056	1	US-08-121-202-1
25	32.8	4.1	7881	2	US-08-751-189-1
26	32.8	4.1	7881	2	US-09-060-836-1
27	32.8	4.1	7881	4	US-09-184-445-1

C 28	32.8	4.1	8241	6	5171844-1	Patent No. 5171844
C 29	32.8	4.1	8967	1	US-08-366-851A-1	Sequence 1, Appli
C 30	32.8	4.1	9009	1	US-07-864-004B-3	Sequence 3, Appli
C 31	32.8	4.1	9009	1	US-08-251-937A-3	Sequence 3, Appli
C 32	32.8	4.1	9009	1	US-08-212-133A-1	Sequence 1, Appli
C 33	32.8	4.1	9009	1	US-08-474-503-1	Sequence 1, Appli
C 34	32.8	4.1	9009	2	US-08-670-707A-1	Sequence 1, Appli
C 35	32.8	4.1	9009	4	US-09-037-601-1	Sequence 1, Appli
C 36	32.8	4.1	9009	4	US-09-315-179-1	Sequence 1, Appli
C 37	32.8	4.1	9009	4	US-09-523-656-1	Sequence 1, Appli
C 38	32.8	4.1	9009	5	PCT-US93-03275-3	Sequence 3, Appli
C 39	32.8	4.1	9009	5	PCT-US94-13200-1	Sequence 1, Appli
C 40	32.4	4.0	630	4	US-08-952-736A-4	Sequence 4, Appli
C 41	32.4	4.0	927	4	US-08-952-736A-1	Sequence 1, Appli
C 42	32.4	4.0	1356	4	US-08-952-736A-2	Sequence 2, Appli
C 43	32.2	4.0	1990	4	US-09-149-476-281	Sequence 281, App
C 44	32.2	4.0	2041	4	US-09-149-476-131	Sequence 131, App
C 45	31.8	4.0	1008	4	US-08-235-836C-130	Sequence 130, App

ALIGNMENTS

RESULT 1

US-09-564-805-28
; Sequence 28, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavrigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564.805
; CURRENT FILING DATE: 2000-05-05
; PRIOR FILING DATE: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 26664
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (910)..(13104)
; OTHER INFORMATION: exon 1: 910-1154; exon 2: 1736-1786; exon 3:
; OTHER INFORMATION: 1925-1995; exon 4: 3025-3089; exon 5: 4361-4418;
; OTHER INFORMATION: exon 6: 5582-5850; exon 7: 7075-7194; exon 8:
; OTHER INFORMATION: 8186-8244; exon 9: 12878-12936; exon 10:
; NAME/KEY: misc feature
; LOCATION: (13756)..(22917)
; OTHER INFORMATION: exon 11: 13756-13868; exon 12: 15283-15378; exon
; OTHER INFORMATION: 13: 16278-16416; exon 14: 16498-16583; exon 15:
; OTHER INFORMATION: 18583-18701; exon 16: 20349-20445; exon 17:
; OTHER INFORMATION: 22172-22310; exon 18: 22879-22917
; NAME/KEY: misc feature
; LOCATION: (23045)..(26452)
; OTHER INFORMATION: exon 19: 23045-23154; exon 20: 23795-23895; exon
; OTHER INFORMATION: 21: 23973-24093; exon 22: 24354-24432; exon 23:
; OTHER INFORMATION: 25026-25170; exon 24: 25812-26036; polyadenylation
; OTHER INFORMATION: signal: 26447-26452
; NAME/KEY: variation
; LOCATION: (826)..(23879)
; OTHER INFORMATION: 8 at positions 826 and 23180 is G or C; Y at
; OTHER INFORMATION: positions 1314, 5568, 7165, 16431, 1857 and 20486
; OTHER INFORMATION: is C or T; n at position 13128 is t or tgat; r at
; OTHER INFORMATION: positions 2221 and 23879 is A or G.

US-09-564-805-28

Query Match 100.0%; Score 800.6; DB 4; Length 26664;
Best Local Similarity 100.0%; Pred. No. 7.6e-232;
Matches 801; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGTGCTGCTGCTGGTATTTTCAACAGAGGCTGTGGCCACAGCAATCTGCTGGTGCAGAT 60
DB 21800 AGTGCTGCTGGTATTTTCAACAGAGGCTGTGGCCACAGCAATCTGCTGGTGCAGAT 21859

QY 61 TCATTGTTAGGACTAAATGCTTTAAAGCCTCCCTATAAACTTTTTTTTTTTTTTTGTATGC 120
DB 21860 TCATTGTTAGGACTAAATGCTTTAAAGCCTCCCTATAAACTTTTTTTTTTTTTTTGTATGC 21919

QY 121 CCAGCCTTTGTGTAAGTCTACTTGAAGGTTTCAGGGTTCCATGATACCTTTTGTCTA 180
DB 21920 CCAGCCTTTGTGTAAGTCTACTTGAAGGTTTCAGGGTTCCATGATACCTTTTGTCTA 21979

QY 181 TAAAGAGGATGACATGTAAATCACCTTTATGTTAAATTAATTTGGCTTTTATATTAG 240
DB 21980 TAAAGAGGATGACATGTAAATCACCTTTATGTTAAATTAATTTGGCTTTTATATTAG 22039

QY 241 CTCCTCAAGCAAGCAGGAGACAGAAATTTCTGAGTTGCTCTTGGTCTCTGTCCAA 300
DB 22040 CTCCTCAAGCAAGCAGGAGACAGAAATTTCTGAGTTGCTCTTGGTCTCTGTCCAA 22099

QY 301 ACCAGACATCAGCCTCTGAAACCATGACAGTCTTCTAGTGGCAGTACTCTCTCTCT 360
DB 22100 ACCAGACATCAGCCTCTGAAACCATGACAGTCTTCTAGTGGCAGTACTCTCTCTCT 22159

QY 361 TCTCTTTGACGCCCGACACGCTCTGCTACTGAGTCTGTGAGTGGGACATTTGGGCA 420
DB 22160 TCTCTTTGACGCCCGACACGCTCTGCTACTGAGTCTGTGAGTGGGACATTTGGGCA 22219

QY 421 GCTGTGCCGTCAATTACGGAGACAGGTGACAGGGTCTTGGGACCCCTGCTGTGTT 480
DB 22220 GCTGTGCCGTCAATTACGGAGACAGGTGACAGGGTCTTGGGACCCCTGCTGTGTT 22279

QY 481 TGTGTCCACCTGACGAGATCACACAGGTGAGTGTGGGCTGGACCAAAAGCTGG 540
DB 22280 TGTGTCCACCTGACGAGATCACACAGGTGAGTGTGGGCTGGACCAAAAGCTGG 22339

QY 541 ASCCTGGAGGCACTGCCAGCTTGAGTTGCCCTTGGCTGCGTCTTTCTCGCTT 600
DB 22340 ASCCTGGAGGCACTGCCAGCTTGAGTTGGCCCTTGGCTGCGTCTTTCTCGCTT 22399

QY 601 CCAAACTTGCCAGAGCTTTTGTACTCTCTGCTAGGAAATGGTTTTTGGCAAAAC 660
DB 22400 CCAAACTTGCCAGAGCTTTTGTACTCTCTGCTAGGAAATGGTTTTTGGCAAAAC 22459

QY 661 TCAACATAGTCTCTTGGCCCAAGAAATGCTTCTCTCTCTGCTGAGTCTCTTTCTCTGC 720
DB 22460 TCAACATAGTCTCTTGGCCCAAGAAATGCTTCTCTCTCTGCTGAGTCTCTTTCTCTGC 22519

QY 721 ASCAGACAGGTTTGAGTTTACCAGCCTTCTGAGTCTTGAATCTCACAGGCTGCT 780
DB 22520 ASCAGACAGGTTTGAGTTTACCAGCCTTCTGAGTCTTGAATCTCACAGGCTGCT 22579

QY 781 CAGCGAAGCTTTGACCGGAT 801
DB 22580 CAGCGAAGCTTTGACCGGAT 22600

RESULT 2

US-09-564-805-1
; Sequence 1, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.

; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2481
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
US-09-564-805-1

Query Match 18.2%; Score 145.8; DB 4; Length 2481;
Best Local Similarity 89.1%; Pred. No. 3.8e-34;
Matches 156; Conservative 1; Mismatches 18; Indels 0; Gaps 0;

QY 338 AGTGGCAGTGACTCTCTTCCTCTCTCTCTGCGAGCCCGGACACGCTCTCTCTACTGGAC 397
DB 1486 ATTCGAATGTCAGTGCACACTTGTCAACATAAGCCCGGACACGCTCTCTCTACTGGAC 1545

QY 398 TGTGCTGAGGCACRTTTGGCAGCTGTGCCGTCAATTACGAGACCCAGGTGGACAGGTC 457
DB 1546 TGTGCTGAGGCACATTTGGGCAGCTGTGCCGTCAATTACGAGACCCAGGTGGACAGGTC 1605

QY 458 CTGGCACCCTGGCTGCTGTGTTGTCTCCACCTGCGACGAGATCACCACCG 512
DB 1606 CTGGCACCCTGGCTGCTGTGTTGTCTCCACCTGCGACGAGATCACCACCG 1660

RESULT 3

US-09-564-805-225
; Sequence 225, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 225
; LENGTH: 2892
; TYPE: DNA
; ORGANISM: Gorilla gorilla
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
US-09-564-805-225

Query Match 18.2%; Score 145.8; DB 4; Length 2892;
Best Local Similarity 89.1%; Pred. No. 4.1e-34;
Matches 156; Conservative 1; Mismatches 18; Indels 0; Gaps 0;

QY 338 AGTGGCAGTGACTCTCTTCCTCTCTCTCTGCGAGCCCGGACACGCTCTCTCTACTGGAC 397
DB 1486 ATTCGAATGTCAGTGCACACTTGTCAACATAAGCCCGGACACGCTCTCTCTACTGGAC 1545

; CURRENT APPLICATION NUMBER: US/09/376,728
; CURRENT FILING DATE: 1999-08-17
; EARLIER APPLICATION NUMBER: US 60/097,242
; EARLIER FILING DATE: 1998-08-20
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 840
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (51)...(623)
US-09-376-728-1

Query Match 4.9%; Score 39.4; DB 4; Length 840;
Best Local Similarity 52.8%; Pred. No. 0.034;
Matches 85; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
QY 307 CATCAGCTCTGAACCATCAGACGCTCTCTGCTACTGACCTGTGAGGACACRTTTGGGAGCTGTG 366
Db 615 CAACAGCGCGGGCGGCTTCTGCT 556
QY 367 CTGACGCCCGCACACGCTCTCTGCTACTGACCTGTGAGGACACRTTTGGGAGCTGTG 426
Db 555 CGCGCGGCTGGCGGCTTCTGCT 496
QY 427 CCGCTATTACGGACACAGGTGACAGGCTCTGGGACACC 467
Db 495 CTTGATGACGGCGGCGGCTGTGCTGTAGGCTCTCTCCAC 455

RESULT 10
US-09-134-001C-1951/c
; Sequence 1951, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1951
; LENGTH: 810
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1951

Query Match 4.6%; Score 36.8; DB 4; Length 810;
Best Local Similarity 69.4%; Pred. No. 0.2;
Matches 50; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
QY 47 CTGATGGTCAGATTCATTGTTAGGACTAAATGCTTTAAGCCTCTCTATAAAGCTTTT 106
Db 139 CAGCAGTTTCAGATTCATTGATGAAGATAAATCCCTAAACATTCCTATAAATGTTTAA 80
QY 107 TTTTCTTTTGTAT 118
Db 79 TTGCTTTCTTAT 68

RESULT 11
US-08-487-826B-13/c
; Sequence 13, Application US/08487826B
; Patent No. 5993827
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.

; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,826B
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelisen, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-487-826B-13
Query Match 4.6%; Score 36.6; DB 2; Length 19124;
Best Local Similarity 53.1%; Pred. No. 1.4;
Matches 78; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
QY 93 ATAAACTTTTTTTTTTTTTTTTTTTTGGATGCCAGCCTTTGTGTAAGTCTACTTCAAGGGTT 152
Db 15941 ATTAATAATTTTTTTTTTTTTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTAAT 15882
QY 153 TCAGGGTTCATGGATCTCTTTCTGTATAAGAGGATGACACATGTAATAACACCTTTTA 212
Db 15881 AAATTTTTTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 15822
QY 213 TGGTTAAATTAATCGCTTTTATATTA 239
Db 15821 TTATTTAAATAAATTTTTTTTTTATTTTA 15795
RESULT 12
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24

```
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match          4.6%; Score 36.6; DB 4; Length 4403765;
Best Local Similarity 50.9%; Pred. No. 25;
Matches 84; Conservative 1; Mismatches 80; Indels 0; Gaps 0;

Qy 359 CTTCTCTTCTGAGCCCGACACGCTCTCTGCTACTGGACTGTGTGAGGGACACRTTTGGG 418
      |||||
Db 435315 CTTGCTTTCGTGGCTCAAAATGGGCTCAGCAACTCGACGAGCGCGCGCGCTCGGC 435374
      |||||

Qy 419 CAGCTGTGCGGTCATTAGGAGACACAGGTGGACAGGGTCTTGGGCACCCCTGCTGCTGTG 478
      |||||
Db 435375 CGCCTCGTCCACACAGGGGTGCGCGGGGTCCGCGCGACCGCAGGTCTCTCAACCGCG 435434
      |||||

Qy 479 TTTGTGTCCCACTGCACGCGAGATCACACACGGTGAGTGTGGG 523
      |||||
Db 435435 CATGTGCGCCAGAACTGCGCGAACCCTGCGGACTGGCCGG 435479
      |||||

RESULT 13
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match          4.6%; Score 36.6; DB 4; Length 4411529;
Best Local Similarity 50.9%; Pred. No. 25;
Matches 84; Conservative 1; Mismatches 80; Indels 0; Gaps 0;

Qy 359 CTTCTCTTCTGAGCCCGACACGCTCTCTGCTACTGGACTGTGTGAGGGACACRTTTGGG 418
      |||||
Db 435233 CTTGCTTTCGTGGCTCAAAATGGGCTCAGCAACTCGACGAGCGCGCGCGCTCGGC 435292
      |||||

Qy 419 CAGCTGTGCGGTCATTAGGAGACACAGGTGGACAGGTCTCTGGGCACCCCTGCTGCTGTG 478
      |||||
Db 435293 CGCCTCGTCCACACGGGTGCGCGGGGTCCGCGCGACCGCAGGTCTCTCAACCGCGG 435352
      |||||

Qy 479 TTTGTGTCCCACTGCACGCGAGATCACACACGGTGAGTGTGGG 523
      |||||
Db 435353 CATGTGCGCCAGAACTGSCCGAACCCTGCGGACTGGCCGG 435397
      |||||

RESULT 14
US-09-461-697-261/c
; Sequence 261, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Puranam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; TITLE OF INVENTION: CELL DEATH
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/461.697
; CURRENT FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 261
; LENGTH: 1182
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-461-697-261

Query Match          4.5%; Score 36; DB 4; Length 1182;
Best Local Similarity 51.9%; Pred. No. 0.44;
Matches 81; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

Qy 99 TTTTCTTTTCTTTTGTGATGCCAGCCTTTGTGTAAGTCTACTTTGAAAGGTTTCAGGG 158
      |||||
Db 1152 TTTTCTTTTCTTTTATTAATACCATCAATTTGAAATTTTACAGGAACCAAGATTCAAG 1093
      |||||

Qy 159 TTCCATGATGACTCTTTTGTCTATAAAGAGATGACACATGTAAATCACCTTTTATGGTTA 218
      |||||
Db 1092 CTCCTTAGTGCTACTGTATTTTATGTTGCACGACAGACGTCATACACAGTTTCAAT 1033
      |||||

Qy 219 AATTAATGGCTTTTATATTAGTCTCTCAAGCAAA 254
      |||||
Db 1032 AGTAATTTTTCATCATATAGGTTTTCCTTAAAAAAA 997
      |||||

RESULT 15
US-09-149-476-255/c
; Sequence 255, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149.476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040.162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040.333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038.621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040.626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040.334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040.336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040.163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047.600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047.615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047.597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047.502
; EARLIER FILING DATE: 1997-05-23
```


GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 15:56:54 ; Search time 681.842 Seconds
(without alignments)
11900.029 Million cell updates/sec

Title: US-09-434-382-28_COPY_26164_26664
Perfect score: 501
Sequence: 1 ggtatggagctgtcgcgagg.....ttcgcgaagtctttttgaca 501

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_nam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length	Description
1	469	93.6 1032 13	BM479686
2	303.2	60.5 345 10	AW407520
3	303	60.5 396 9	AA811170
4	302.6	60.4 372 9	AI500718
5	301.4	60.2 478 10	AW592601
6	301.4	60.2 499 10	AW510825

C	7	301.4	60.2	531	10	AW572950	hfl7h05.x
C	8	301.4	60.2	692	14	BQ181829	UI-H-EUO-
C	9	301.4	60.2	712	14	BM977567	UI-CF-EN1
C	10	299.8	59.8	394	10	AW248468	AW20640.3
C	11	299.4	59.8	527	9	AI089646	qb16g07.x
C	12	298.4	59.6	316	10	AW511765	AW511765 xu76f03.x
C	13	298.2	59.5	318	9	AA766184	calaf09.s
C	14	298.2	59.5	410	9	AI033108	ow98g08.s
C	15	297.2	59.3	374	9	AA223338	zr05h05.s
C	16	296.6	59.2	354	14	BQ030358	UI-H-DT0-
C	17	296.6	59.2	390	9	AA605004	no47g05.s
C	18	296.6	59.2	406	9	AA583412	nn42b10.s
C	19	296.6	59.2	416	9	AI141263	q46h05.s
C	20	296.6	59.2	433	9	AI357786	qu98d07.x
C	21	296.6	59.2	461	10	AW592223	hf41a01.x
C	22	296.6	59.2	479	9	AA716607	zg68g07.s
C	23	296.6	59.2	489	9	AI803400	tc42f03.x
C	24	296.6	59.2	489	9	AA534478	nf76f10.s
C	25	296.6	59.2	531	9	AI937465	wp77e01.x
C	26	296.6	59.2	536	14	BM681558	UI-E-EJ0-
C	27	296.6	59.2	614	13	BI870047	603393914
C	28	296.2	59.1	491	9	AA676661	zj67h01.s
C	29	295.6	59.0	477	9	AI804749	tu42d02.x
C	30	294.6	58.8	409	12	BF477438	nac60r05.
C	31	294.6	58.8	439	9	AI201492	qs74b03.x
C	32	294.6	58.8	494	12	BE858252	7g21a09.x
C	33	294.6	58.8	535	9	AI158921	AU158921
C	34	294.6	58.8	536	9	AI153280	AU153280
C	35	293.6	58.6	372	14	D29176	HUMNK212.Hu
C	36	293.6	58.6	481	9	AI149928	AU149928
C	37	293.6	58.6	541	9	AI148489	AU148489
C	38	293.4	58.6	306	9	AA862956	cg98e07.s
C	39	291.8	58.2	409	9	AI831155	wj81f07.x
C	40	291.6	58.2	352	14	R44712	yg25h08.s1
C	41	290.2	57.9	488	13	BM353522	ig56c05.x
C	42	289.6	57.8	329	14	R58845	NIB2016.NoI
C	43	289.4	57.8	398	9	AA948352	oq46d06.s
C	44	288.6	57.6	414	9	AA134386	zo24d10.s
C	45	287.2	57.3	367	14	BQ549604	ik88h10.x

ALIGNMENTS

RESULT 1
BM479686
LOCUS BM479686 1032 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6418543 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5502530
5', mRNA sequence.
ACCESSION BM479686
VERSION BM479686.1 GI:18528728
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1032)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12140 row: n column: 03
High quality sequence stop: 720.
Location/Qualifiers 1. .1032
FEATURES
source

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5502530"
/clone_lib="NIH_MGC_67"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."
BASE COUNT      253 a      234 c      295 g      247 t      3 others
ORIGIN
Query Match      93.6%; Score 469; DB 13; Length 1032;
Best Local Similarity 97.6%; Pred. No. 2.8e-131;
Matches 489; Conservative 0; Mismatches 5; Indels 7; Gaps 1;
QY 1 GGTATGGAGCTGTGCGGAGGCTTGGCTCCCATAAAGCACTAGTCTATAGATGCTCTT 60
Db 170 GGTATGGAGCTGTGCGGAGGCTTGGGTTCCCATAAAGCACTAGTCTATAGATGCTCTT 229
QY 61 AGGACTGTGCTGCGACAGCGCGGCGAGAGGCTGCCACACGGAAGCAAGCAGATGA 120
Db 230 AGGACTGTGCTGCGACAGCTGCGGCGAGAGGCTGCCACACGGAAGCAAGCAGATGA 289
QY 121 ACTAATTTTCATTTCAAGCAGTCTTTTAAAGAGTCTTGGAAACAGACGCGGCACCTTTC 180
Db 290 ACTAATTTTCATTTCAAGCAGTCTTTTAAAGAGTCTTGGAAACAGACGCGGCACCTTTC 349
QY 181 CTCTAATCAGCAAAAGTGATTCCTCTGCACACGAGACAGCAAGCAGATGACAGT 240
Db 350 CTCTAATCAGCAAAAGTGATTCCTCTGCACACGAGACAGCAAGCAGATGACAGT 409
QY 241 GGTCTAAGTGTCCGAGACTTAAAGAAATAGTATTTACGCTGCAATAAAGATGAGTTTG 300
Db 410 GGTCTAAGTGTCCGAGACTTAAAGAAATAGTATTTACGCTGCAATAAAGATGAGTTTG 469
QY 301 CAATTTGTAGTCTTTTGGTCTCTCTGCTGCTACAGACAGAGGCTGCTGTGACAC 360
Db 470 CAATTTGTAGTCTTTTGGTCTCTCTGCTGCTACAGACAGAGGCTGCTGTGACAC 522
QY 361 ACCTTTGGAGAGGCTCTCTGCTGCTAGTGTGGAGCTGCTGTGACCGGCTGCTGG 420
Db 523 ACCTTTGGAGAGGCTCTCTGCTGCTAGTGTGGAGCTGCTGTGACCGGCTGCTGG 582
QY 421 AAGAAAGTCAAGTCCCGTCTGTAGTGAACCTCTGGAACCTGCTCTCAGAGACCCCTT 480
Db 583 AAGAAAGTCAAGTCCCGTCTGTAGTGAACCTCTGGAACCTGCTCTCAGAGACCCCTT 642
QY 481 ATTCGCCAAGTCTTTTGACA 501
Db 643 ATTCGCCAAGTCTTTTGACA 663
```

```
RESULT 2
AW407520/c      AW407520      345 bp      mRNA      linear      EST 16-FEB-2000
LOCUS      UI-HF-BMO-adl-b-07-0-UI.r1 NIH_MGC_38 Homo sapiens cDNA clone
DEFINITION      IMAGE:3061957 5', mRNA sequence.
ACCESSION      AW407520
VERSION      AW407520.1 GI:6926577
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 345)
REFERENCE      NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS      National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE      Unpublished (1999)
JOURNAL
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Eco RI site shown at the beginning of the sequence.
```

```
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: M.B. Soares Lab
cDNA Library Arrayed by: M.B. Soares Lab
DNA sequencing by: M.B. Soares Lab
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: M13 Forward
Location/Qualifiers
1. .345
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3061957"
/clone_lib="NIH_MGC_38"
/tissue_type="lymph"
/cell_type="germinal center B cells"
/lab_host="DH10B (LT1)"
/notes="Vector: pT7T3-Pac; Site 1: NotI; Site 2: Eco RI;
Constructed from size fractionated cytoplasmic mRNA
(2.5-3.5kb). Directionally cloned. Cells provided by Louis
M. Staudt, Ph.D. Library preparation by Maria de Fatima
Bonaldo, Ph.D. and M. Bento Soares, Ph.D."
BASE COUNT      73 a      94 c      82 g      96 t
ORIGIN
Query Match      60.5%; Score 303.2; DB 10; Length 345;
Best Local Similarity 99.0%; Pred. No. 3.7e-81;
Matches 305; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 GGTATGAGCTGTGCGGAGGCTTGGCTCCCATAAAGCACTAGTCTATAGATGCTCTT 60
Db 308 GGTATGAGCTGTGCGGAGGCTTGGCTCCCATAAAGCACTAGTCTATAGATGCTCTT 249
QY 61 AGGACTGTGCTGCGACAGCGCGGCGAGGCTGCCACACGGAAGCAAGCAGATGA 120
Db 248 AGGACTGTGCTGCGACAGCGCGGCGAGGCTGCCACACGGAAGCAAGCAGATGA 189
QY 121 ACTAATTTTCATTTCAAGCAGTCTTTTAAAGAGTCTTGGAAACAGACGCGGCACCTTTC 180
Db 188 ACTAATTTTCATTTCAAGCAGTCTTTTAAAGAGTCTTGGAAACAGACGCGGCACCTTTC 129
QY 181 CTCTAATCAGCAAAAGTGATTCCTCTGCACACGAGACAGCAGTAAACAGGATCAGTG 240
Db 128 CTCTAATCAGCAAAAGTGATTCCTCTGCACACGAGACAGCAGTAAACAGGATCAGTG 69
QY 241 GGTCTAAGTGTCCGAGACTTAAAGAAATAGTATTTTCAGCTGCAATAAAGATGAGTTTG 300
Db 68 GGTCTAAGTGTCCGAGACTTAAAGAAATAGTATTTTCAGCTGCAATAAAGATGAGTTTG 9
QY 301 CAATTTGTG 308
Db 8 CCCTCGTG 1
```

```
RESULT 3
AA811170/c      AA811170      396 bp      mRNA      linear      EST 19-FEB-1998
LOCUS      ob42c03.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1334020 3',
DEFINITION      mRNA sequence.
ACCESSION      AA811170
VERSION      AA811170.1 GI:2880781
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 396)
REFERENCE      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE      Tumor Gene Index
JOURNAL      Unpublished (1997)
COMMENT      Contact: Robert Strausberg, Ph.D.
```

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman, Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bbrp/image/image.html
Insert Length: 885 Std Error: 0.00
Seg primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 377.

FEATURES

Location/Qualifiers
1..396
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1334020"
/clone_lib="NCI-CGAP_GCBI"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr. David Allman (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGCGCTCATTTTTTTTTTTT-3'
1. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."
81 a 102 c 97 g 116 t

BASE COUNT
ORIGIN

Query Match 60.5%; Score 303; DB 9; Length 396;
Best Local Similarity 100.0%; Pred. No. 4.7e-81;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGTATGGAGCTGTCCGAGGCTTGGGCTCCACATAGCACTAGTCTATAGATGCTCTT 60
Db |||||
315 GGTATGGAGCTGTCCGAGGCTTGGGCTCCACATAGCACTAGTCTATAGATGCTCTT 256
Qy 61 AGGACTGTGTCCGACAGCGCGGCGGCGGCGGCTGCCACACGGAAGCAGCATCA 120
Db |||||
255 AGGACTGTGTCCGACAGCGCGGCGGCGGCGGCTGCCACACGGAAGCAGCATCA 196
Qy 121 ACTAATTTTCATTTCAAGGAGTCTTTTAAAGAGTCTTTGGAACAGACGCGGCACCTTTC 180
Db |||||
195 ACTAATTTTCATTTCAAGGAGTCTTTTAAAGAGTCTTTGGAACAGACGCGGCACCTTTC 136
Qy 181 CTCTAATCCAGCAAGTATTCCTCCACACGACAGCAGCAGAGTAAAGATCAGTG 240
Db |||||
135 CTCTAATCCAGCAAGTATTCCTCCACACGACAGCAGCAGAGTAAAGATCAGTG 76
Qy 241 GGTCTAAGTGTCCGAGACTTAAAGAAAATAGTATTTTCAGTGCATTAAGATTTGATTTG 300
Db |||||
75 GGTCTAAGTGTCCGAGACTTAAAGAAAATAGTATTTTCAGTGCATTAAGATTTGATTTG 16
Qy 301 CAA 303
Db |||||
15 CAA 13

RESULT 4
AI500718/c
LOCUS AI500718 372 bp mRNA linear EST 14-APR-1999
DEFINITION tn94b10.x1 NCI-CGAP_Ut2 Homo sapiens cDNA clone IMAGE:217179.3', mRNA sequence.
ACCESSION AI500718
VERSION AI500718.1 GI:4392700

KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM human.
REFERENCE 1 (bases 1 to 372)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Greg Lennon, Ph.D.
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bbrp/image/image.html
Insert Length: 737 Std Error: 0.00
Seg primer: -40UP from Gibco
High quality sequence stop: 342.

FEATURES

source
1..372
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:217179"
/clone_lib="NCI-CGAP_Ut2"
/tissue_type="moderately-differentiated endometrial adenocarcinoma, 3 pooled tumors"
/lab_host="DH10B"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.85 kb. Life Technologies catalog #: 11539-012"
80 a 100 c 82 g 110 t

BASE COUNT
ORIGIN

Query Match 60.4%; Score 302.6; DB 9; Length 372;
Best Local Similarity 98.7%; Pred. No. 6e-81;
Matches 305; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 GGTATGGAGCTGTCCGAGGCTTGGGCTCCACATAGCACTAGTCTATAGATGCTCTT 60
Db |||||
315 GGTATGGAGCTGTCCGAGGCTTGGGCTCCACATAGCACTAGTCTATAGATGCTCTT 256
Qy 61 AGGACTGTGTCCGACAGCGCGGCGGCGGCGGCTGCCACACGGAAGCAGCATCA 120
Db |||||
255 AGGACTGTGTCCGACAGCGCGGCGGCGGCGGCTGCCACACGGAAGCAGCATCA 196
Qy 121 ACTAATTTTCATTTCAAGGAGTCTTTTAAAGAGTCTTTGGAACAGACGCGGCACCTTTC 180
Db |||||
195 ACTAATTTTCATTTCAAGGAGTCTTTTAAAGAGTCTTTGGAACAGACGCGGCACCTTTC 136
Qy 181 CTCTAATCCAGCAAGTATTCCTCCGACACGACAGCAGAGTAAAGATCAGTG 240
Db |||||
135 CTCTAATCCAGCAAGTATTCCTCCGACACGACAGCAGAGTAAAGATCAGTG 76
Qy 241 GGTCTAAGTGTCCGAGACTTAAAGAAAATAGTATTTTCAGTGCATTAAGATTTGATTTG 300
Db |||||
75 GGTCTAAGTGTCCGAGACTTAAAGAAAATAGTATTTTCAGTGCATTAAGATTTGATTTG 16
Qy 301 CAATTGTGA 309
Db |||||
15 CAATTGTGA 7

RESULT 5
AW592601/c
LOCUS AW592601 478 bp mRNA linear EST 22-MAR-2000
DEFINITION hf45a09.x1 Soares NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:2934808.3', mRNA sequence.

```

ACCESSION      AW592601
VERSION        AW592601.1  GI:7279786
KEYWORDS       EST.
SOURCE         human.
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 478)
AUTHORS        NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE          National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
               Tumor Gene Index
JOURNAL         Unpublished (1997)
COMMENT        Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-remail.nih.gov
               This clone is available royalty-free through LLNL ; contact the
               IMAGE Consortium (info@image.llnl.gov) for further information.
               Seq primer: -40UP from Gibco
               High quality sequence stop: 451.
               Location/Qualifiers
FEATURES       source
               1..478
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="IMAGE:2934808"
               /clone_lib="Soares_NFL_T_GBC_S1"
               /lab_host="DH10B"
               /note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
               a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
               Equal amounts of plasmid DNA from three normalized
               libraries (fetal lung NBHL19W, testis NHT, and B-cell
               NCI CGAP GCB1) were mixed, and ss circles were made in
               vitro. Following HAP purification, this DNA was used as
               tracer in a subtractive hybridization reaction. The driver
               was PCR-amplified cDNAs from pools of 5,000 clones made
               from the same 3 libraries. The pools consisted of
               I.M.A.G.E. clones 297480-302087, 682632-687239,
               726408-728711, and 729096-731399. Subtraction by Bento
               Soares and M. Fatima Bonaldo."
BASE COUNT     92 a 132 c 122 g 132 t
ORIGIN
Query Match    60.2%; Score 301.4; DB 10; Length 478;
Best Local Similarity 99.7%; Pred. No. 1.6e-80;
Matches 302; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GGTATGGAGCTGTGCGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCCTCTT 60
Db 305 GGTATGGAGCTGTGCGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCCTCTT 246
Qy 61 AGGACTGGTGCTGGCAGCGCGGCGGCGGAGGCTGCCACAGCAAGCAGATGA 120
Db 245 AGGACTGGTGCTGGCAGCGCGGCGGCGGAGGCTGCCACAGCAAGCAGATGA 186
Qy 121 ACTAATTTTCATTTCAAGGCAGTCTTTTAAAGAAAGTCTTTGAAACAGACGCGGCACCTTTC 180
Db 185 ACTAATTTTCATTTCAAGGCAGTCTTTTAAAGAAAGTCTTTGAAACAGACGCGGCACCTTTC 126
Qy 181 CTCCTAATCCAGCAAAAGTATTCCTTCGACACAGAGCAAGCAGAGTAACAGATCAGTG 240
Db 125 CTCCTAATCCAGCAAAAGTATTCCTTCGACACAGAGCAAGCAGAGTAACAGATCAGTG 66
Qy 241 GGTCTAAGTGTCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAAGATTGAGTTTG 300
Db 65 GGTCTAAGTGTCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAAGATTGAGTTTG 6
Qy 301 CAA 303
Db 5 CAA 3
RESULT 6
AW510825/c
LOCUS         hd40b11.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
DEFINITION
IMAGE:2911965 3', mRNA sequence.
ACCESSION      AW510825
VERSION        AW510825.1  GI:7148903
KEYWORDS       EST.
SOURCE         human.
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 499)
AUTHORS        NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE          National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
               Tumor Gene Index
JOURNAL         Unpublished (1997)
COMMENT        Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-remail.nih.gov
               This clone is available royalty-free through LLNL ; contact the
               IMAGE Consortium (info@image.llnl.gov) for further information.
               Seq primer: -40UP from Gibco
               High quality sequence stop: 470.
               Location/Qualifiers
FEATURES       source
               1..499
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="IMAGE:2911965"
               /clone_lib="Soares_NFL_T_GBC_S1"
               /lab_host="DH10B"
               /note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
               a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
               Equal amounts of plasmid DNA from three normalized
               libraries (fetal lung NBHL19W, testis NHT, and B-cell
               NCI CGAP GCB1) were mixed, and ss circles were made in
               vitro. Following HAP purification, this DNA was used as
               tracer in a subtractive hybridization reaction. The driver
               was PCR-amplified cDNAs from pools of 5,000 clones made
               from the same 3 libraries. The pools consisted of
               I.M.A.G.E. clones 297480-302087, 682632-687239,
               726408-728711, and 729096-731399. Subtraction by Bento
               Soares and M. Fatima Bonaldo."
BASE COUNT     94 a 142 c 126 g 137 t
ORIGIN
Query Match    60.2%; Score 301.4; DB 10; Length 499;
Best Local Similarity 99.7%; Pred. No. 1.7e-80;
Matches 302; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GGTATGGAGCTGTGCGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCCTCTT 60
Db 303 GGTATGGAGCTGTGCGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCCTCTT 244
Qy 61 AGGACTGGTGCTGGCAGCGCGGCGGCGGAGGCTGCCACAGCAAGCAGATGA 120
Db 243 AGGACTGGTGCTGGCAGCGCGGCGGCGGAGGCTGCCACAGCAAGCAGATGA 184
Qy 121 ACTAATTTTCATTTCAAGGCAGTCTTTTAAAGAAAGTCTTTGAAACAGACGCGGCACCTTTC 180
Db 183 ACTAATTTTCATTTCAAGGCAGTCTTTTAAAGAAAGTCTTTGAAACAGACGCGGCACCTTTC 124
Qy 181 CTCCTAATCCAGCAAAAGTATTCCTTCGACACAGAGCAAGCAGAGTAACAGATCAGTG 240
Db 123 CTCCTAATCCAGCAAAAGTATTCCTTCGACACAGAGCAAGCAGAGTAACAGATCAGTG 64
Qy 241 GGTCTAAGTGTCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAAGATTGAGTTTG 300
Db 63 GGTCTAAGTGTCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAAGATTGAGTTTG 4
Qy 301 CAA 303
Db 3 CAA 1
RESULT 7
AW572950/c
LOCUS         AW572950
DEFINITION    531 bp mRNA linear EST 13-MAR-2000

```



```

DEFINITION hfl7h05.x1 Soares NFL_T_GBC_S1 Homo sapiens cDNA clone
IMAGE:2932185 3', mRNA sequence.
ACCESSION AW572950
VERSION AW572950.1 GI:7237683
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 531)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco
High quality sequence stop: 459.
FEATURES
    source
        1..531
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:2932185"
        /clone_lib="Soares_NFL_T_GBC_S1"
        /lab_host="DH10B"
        /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
        a modified polylinker; Site 1: Not I; Site 2: Eco RI;
        Equal amounts of plasmid DNA from three normalized
        libraries (fetal lung NBHL19W, testis NHR, and B-cell
        NCI-CGAP GCB1) were mixed, and ss circles were made in
        vitro. Following HAP purification, this DNA was used as
        tracer in a subtractive hybridization reaction. The driver
        was PCR-amplified cDNAs from pools of 5,000 clones made
        from the same 3 libraries. The pools consisted of
        I.M.A.G.E. clones 297480-302087, 682632-687239,
        726408-728711, and 729096-731399. Subtraction by Bento
        Soares and M. Fatima Bonaldo. "
BASE COUNT 99 a 154 c 136 g 142 t
ORIGIN
Query Match 60.2%; Score 301.4; DB 10; Length 531;
Best Local Similarity 99.7%; Pred. No. 1.8e-80;
Matches 302; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGTATGGAGCTGTGCCAGGCTTGGGCTCCCATAGCACTAGTCTATAGATGCTCTT 60
Db 305 GGTATGGAGCTGTGCCAGGCTTGGGCTCCCATAGCACTAGTCTATAGATGCTCTT 246
QY 61 AGGACTGGTGGCCAGCCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA 120
Db 245 AGGACTGGTGGCCAGCCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA 186
QY 121 ACTAATTTTCATTCAAGGAGTTTTTAAAGAGTCTTTGGAACAGACGCGGCACCTTTC 180
Db 185 ACTAATTTTCATTCAAGGAGTTTTTAAAGAGTCTTTGGAACAGACGCGGCACCTTTC 126
QY 181 CTCTAATCCAGCAAGTGTATCCCTGCACACGAGACGAGGATACAGGATCAGTG 240
Db 125 CTCTAATCCAGCAAGTGTATCCCTGCACACGAGACGAGGATACAGGATCAGTG 66
QY 241 GGTCTAAGTGTCCGAGACCTTAACGAAAATAGTATTTTCAGCTGCAATAAAGATTGTTTG 300
Db 65 GGTCTAAGTGTCCGAGACCTTAACGAAAATAGTATTTTCAGCTGCAATAAAGATTGTTTG 6
QY 301 CAA 303
Db 5 CAA 3
RESULT 8
BQ181829/c

```

```

LOCUS BQ181829
DEFINITION UI-H-EUO-azv-m-08-0-UI.s1 NCI_CGAP_Carl Homo sapiens cDNA clone
IMAGE: 5854255 3', mRNA sequence.
ACCESSION BQ181829
VERSION BQ181829.1 GI:20357321
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 692)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Seq primer: M13 FORWARD
POLYA=yes.
FEATURES
    Location/Qualifiers
        1..692
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE: 5854255"
        /clone_lib="NCI_CGAP_Carl"
        /tissue_type="Osteoarthritic Cartilage"
        /dev_stage="Adult"
        /lab_host="DH10B (Life Technologies)"
        /note="Organ: Knee; Vector: pT7T3-Pac (Pharmacia) with a
        modified polylinker; Site 1: EcoR I; Site 2: Not I;
        NCI_CGAP_Carl is a cDNA library containing the following
        tissue(s): Osteoarthritic Cartilage The library was
        constructed according to Bonaldo, Lennon and Soares,
        Genome Research, 6:791-806, 1996. First strand cDNA
        synthesis was primed with an oligo-dT primer containing a
        Not I site. Double stranded cDNA was ligated to an EcoR I
        adaptor, digested with Not I, and cloned directionally
        into pT7T3-Pac vector. The oligonucleotide used to prime
        the synthesis of first-strand cDNA contains a library tag
        sequence that is located between the Not I site and the
        (dT)18 tail. The sequence tag for this library is
        TGATCAGCTT.
        TAG_L1B=UI-H-EUO
        TAG_TISSUE=osteoarthritic cartilage
        TAG_SEQ=TGATCAGCTT"
BASE COUNT 127 a 204 c 174 g 187 t
ORIGIN
Query Match 60.2%; Score 301.4; DB 14; Length 692;
Best Local Similarity 99.7%; Pred. No. 2.1e-80;
Matches 302; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGTATGGAGCTGTGCCAGGCTTGGGCTCCCATAGCACTAGTCTATAGATGCTCTT 60
Db 319 GGTATGGAGCTGTGCCAGGCTTGGGCTCCCATAGCACTAGTCTATAGATGCTCTT 260
QY 61 AGGACTGGTGGCCAGCCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA 120
Db 259 AGGACTGGTGGCCAGCCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA 200
QY 121 ACTAATTTTCATTCAAGGAGTTTTTAAAGAGTCTTTGGAACAGACGCGGCACCTTTC 180
Db 199 ACTAATTTTCATTCAAGGAGTTTTTAAAGAGTCTTTGGAACAGACGCGGCACCTTTC 140
QY 181 CTCTAATCCAGCAAGTGTATCCCTGCACACGAGACGAGGATACAGGATCAGTG 240
Db 139 CTCTAATCCAGCAAGTGTATCCCTGCACACGAGACGAGGATACAGGATCAGTG 80

```

```

Qy 241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTCAGCTCAATAAAGATTGAGTTTG 300
|||||
Db 79 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTCAGCTCAATAAAGATTGAGTTTG 20

Qy 301 CAA 303
|||
Db 19 CAA 17

RESULT 9
LOCUS BM977567/c
DEFINITION UI-CF-EN1-aef-c-19-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone
UI-CF-EN1-aef-c-19-0-UI 3', mRNA sequence.
ACCESSION BM977567
VERSION BM977567.1 GI:19596113
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 712)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 9704477
COMMENT Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES             source
    Location/Qualifiers
        1..712
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="UI-CF-EN1-aef-c-19-0-UI"
            /clone_lib="UI-CF-EN1"
            /tissue_type="Primary Lung Cystic Fibrosis Epithelial
            Cells"
            /dev_stages="Adult"
            /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
            /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
            modified polylinker; Site 1: EcoR I; Site 2: Not I;
            UI-CF-EN1 is a normalized cDNA library containing the
            following tissue(s): Primary Lung Cystic Fibrosis
            Epithelial Cells. The library was constructed according to
            Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
            1996. First strand cDNA synthesis was primed with an
            oligo-dT primer containing a Not I site. Double stranded
            cDNA was ligated to an EcoR I adaptor, digested with Not
            I, and cloned directionally into pT7T3-Pac vector. The
            oligonucleotide used to prime the synthesis of
            first-strand cDNA contains a library tag sequence that is
            located between the Not I site and the (dT)18 tail. The
            sequence tag for this library is CTGCTCAGGT.
            TAG LIB=UI-CF-EN1
            TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS
            6hr to LPS 24h
            TAG_SEQ=CTGCTCAGGT"

BASE COUNT          130 a      211 c      176 g      194 t      1 others
ORIGIN

```

```

Query Match          60.2%; Score 301.4; DB 14; Length 712;
Best Local Similarity 99.7%; Pred. No. 2.2e-80;
Matches 302; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCCTCTT 60
|||||
Db 319 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCCTCTT 260

Qy 61 AGGACTGGTGGCTGGCACAGCCGCGGCGGAGGCTGCCACACGGGAAGCAAGCAGATGA 120
|||||
Db 259 AGGACTGGTGGCTGGCACAGCCGCGGCGGAGGCTGCCACACGGGAAGCAAGCAGATGA 200

Qy 121 ACTAATTTCAATTTCAAGCGAGTGTTTTAAAGAAAGTCTTTGGAAACAGACGGCGGCACCTTTTC 180
|||||
Db 199 ACTAATTTCAATTTCAAGCGAGTGTTTTAAAGAAAGTCTTTGGAAACAGACGGCGGCACCTTTTC 140

Qy 181 CTCTAATCCAGCAAAAGTGATTCCTCGCACACAGACAGCAAGCAGAGTAACAGGATCACTG 240
|||||
Db 139 CTCTAATCCAGCAAAAGTGATTCCTCGCACACAGACAGCAAGCAGAGTAACAGGATCACTG 80

Qy 241 GGTCTAAGTGTCCGAGACTTAAACGAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTG 300
|||||
Db 79 GGTCTAAGTGTCCGAGACTTAAACGAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTG 20

Qy 301 CAA 303
|||
Db 19 CAA 17

RESULT 10
LOCUS AW248468/c
DEFINITION 2820640.3prime NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2820640 3',
mRNA sequence.
ACCESSION AW248468
VERSION AW248468.1 GI:6591461
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 394)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Other ESTs: 2820640.5prime
Contact: Robert Strausberg, Ph.D.
Email: cgabs-x@mail.nih.gov
Tissue Procurement: DCTD/DTP cDNA Library Preparation: Ling
Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E.
Consortium (LLNL) DNA Sequencing by: Berkeley MGC sequencing
project Clone distribution: MGC clone distribution information can
be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html Base Calling / Quality
Scores: PHRED from University of Washington Genome Center
Trimming: cross_match from University of Washington Genome Center
PHRAP suite. Poly-T Identification: patmatch.pl from Berkeley
Drosophila Genome Project. University of Washington Genome Center:
http://www.genome.washington.edu Polyadenylation: Based upon the
presence of a XhoI site followed by a run of 14 or more T residues
at the beginning of the sequence, this cDNA insert was
polyadenylated.
Plate: LLCMA row: L column: 17
High quality sequence stop: 213.
Location/Qualifiers
    1..394
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:2820640"
        /clone_lib="NIH_MGC_7"
        /tissue_type="small cell carcinoma"
        /cell_line="MGC3"

```

```

/lab_host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GGCACGAG(G). Size-selected >50bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

Query Match	59.8%;	Score 299.8;	DB 10;	Length 394;
Best Local Similarity	99.3%;	Pred. No. 4.4e-80;		
Matches 301;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
Qy	1	GGTATGGAGCTGTC	CGGAGGCTTGGGCTCC	CACATAAAGCACTAGTCTATAGATGCCTCTT 60
Db	311	GGTATGGAGCTGTC	CGGAGGCTTGGGCTCC	CACATAAAGCACTAGTCTATAGATGCCTCTT 252
Qy	61	AGGACTGTGTGCTGG	CACAGCCGGGGCCAG	GAGGCTGCCACACGGAAGCAAGCAGATGA 120
Db	251	AGGACTGTGTGCTGG	CACAGCCGGGGCCAG	GAGGCTGCCACACGGAAGCAAGCAGATGA 192
Qy	121	ACTAATTTTCATTCA	AGGCGAGTTTTAAAG	AAGTCTTGAAACACAGACGGCGGCACCTTTC 180
Db	191	ACTAATTTTCATTCA	AGGCGAGTTTTAAAG	AAGTCTTGAAACACAGACGGCGGCACCTTTC 132
Qy	181	CTCTAATCCAGCAAG	TGATTTCCCTGCGAC	CACGAGACCAAGCAGAGTAAACAGGATCAGTG 240
Db	131	CTCTAATCCAGCAAG	TGATTTCCCTGCGAC	CACGAGACCAAGCAGAGTAAACAGGATCAGTG 72
Qy	241	GGTCTAAGTGTCCGAG	ACTTAACGAAAATAG	TATTTTCAGCTGCAATAAAGATTTGATTTTG 300
Db	71	GGTCTAAGTGTCCGAG	ACTTAACGAAAATAG	TATTTTCAGCTGCAATAAAGATTTGATTTTG 12
Qy	301	CAA	303	
Db	11	CAA	9	

RESULT 11
AI089646/c
LOCUS
DEFINITION
AI089646 linear mRNA EST 18-AUG-1998
ql16007.x1 Soares_pregnant_uterus_NbHPU Homo sapiens cDNA clone
IMAGE:1696476 3', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
AI089646
AI089646.1 GI:3428705
EST.
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 527)
REFERENCE
AUTHORS
TITLE
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40ml3 fwd. Et' from Amersham
High quality sequence stop: 463.
Location/Qualifiers
1. 527
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1696476"
/clone_lib="Soares_pregnant_uterus_NbHPU"
/sex="Female"
/dev_stage="adult"
/lab_host="DH10B"

```

/notes="Organ: uterus; Vector: pT7T3-Pac; Site_1: Not I;
Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5'
AACTGCAAGAAATTCGGCGCGCCCTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by M. Fatima Bonaldo."
99 a 153 c 136 g 138 t 1 others
BASE COUNT
ORIGIN

```

Query Match	59.8%;	Score	299.4;	DB	9;	Length	527;
Best Local Similarity	99.7%;	Pred. No.	7.1e-80;				
Matches	300;	Conservative	0;	Mismatches	1;	Indels	0;
Gaps	0;						

Qy	1	GGTATGAGCTGTGCCAGAGCTTGGGCTCCACATAGCACTAGTCTTATAGATGCCCTCTT	60
Db	301	GGTATGAGCTGTGCCAGAGCTTGGGCTCCACATAAGCACTAGTCTTATAGATGCCCTCTT	242
Qy	61	AGGACTGGTCTGGCACAGCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA	120
Db	241	AGGACTGGTCTGGCACAGCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA	182
Qy	121	ACTAATTTCAATTTCAAGGCAGTTTTTTAAAGAAGTCTTTGGAAACACAGCGGGCACCTTTTC	180
Db	181	ACTAATTTCAATTTCAAGGCAGTTTTTTAAAGAAGTCTTTGGAAACACAGCGGGCACCTTTTC	122
Qy	181	CTCTAATCCAGCAAGTGATTCCTGTCACACACGAGACAAGCAGAGTAAACAGGATCAGTG	240
Db	121	CTCTAATCCAGCAAGTGATTCCTGTCACACACGAGACAAGCAGAGTAAACAGGATCAGTG	62
Qy	241	GGTCTAAGTGTCCGAGACTTTAAACGAAATAGTATTTTCAGCTTGCAATAAAGATTTCAGTTTG	300
Db	61	GGTCTAAGTGTCCGAGACTTTAAACGAAATAGTATTTTCAGCTTGCAATAAAGATTTCAGTTTG	2
Qy	301	C 301	
Db	1	C 1	

RESULT 12
AWS11765/c
LOCUS
DEFINITION
AWS11765 316 bp mRNA linear EST 03-MAR-2000
xu76f03.x1 NCI_CGAP_Kid8 Homo sapiens cDNA clone IMAGE:2807645 3',
mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Features
source

316 bp mRNA linear EST 03-MAR-2000
xu76f03.x1 NCI_CGAP_Kid8 Homo sapiens cDNA clone IMAGE:2807645 3',
mRNA sequence.
AWS11765
AWS11765.1 GI:7149843
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 316)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapsb@email.nih.gov
Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/iresources.shtml
Seq primer: -40UP from Gibco
High quality sequence stop: 239.
Location/Qualifiers
1. .316
/organism="Homo sapiens"

```

/db_xref="taxon:9606"
/clone="IMAGE:2807645"
/clone_lib="NCI_CGAP_Kid8"
/tissue_type="renal cell tumor"
/lab_host="DH10B"
/note="Organ: kidney; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.2 kb. Life Technologies catalog #:
11524-014"
BASE COUNT      69 a      84 c      72 g      90 t      1 others
ORIGIN

Query Match      59.6%; Score 298.4; DB 10; Length 316;
Best Local Similarity 99.3%; Pred. No. 1e-79;
Matches 299; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGTATGGAGCTGTCGAGGCTTGGGCTCCACATAAGCAGTCTATAGATGCTCTT 60
Db 301 GGTATGGAGCTGTCGAGGCTTGGGCTCCACATNAGCAGTCTATAGATGCTCTT 242

Qy 61 AGGACTGTGCTGTCGACAGCCGCGGCGGAGGCTGCCACACGGAAGCAGATGA 120
Db 241 AGGACTGTGCTGTCGACAGCCGCGGCGGAGGCTGCCACACGGAAGCAGATGA 182

Qy 121 ACTAATTTTCATTTCAAGCAGTTTTTAAAGAGTCTTGGAAACAGACGCGGCACCTTTC 180
Db 181 ACTAATTTTCATTTCAAGCAGTTTTTAAAGAGTCTTGGAAACAGACGCGGCACCTTTC 122

Qy 181 CTCTAATCCAGCAAGTAGTTCCTGCACACGAGACAGCAGAGTAACAGGATCAGTG 240
Db 121 CTCTAATCCAGCAAGTAGTTCCTGCACACGAGACAGCAGAGTAACAGGATCAGTG 62

Qy 241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTACGCTGCAATAAGATTGATTG 300
Db 61 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTACGCTGCAATAAGATTGATTG 2

Qy 301 C 301
Db 1 C 1

RESULT 13
LOCUS AA766184
DEFINITION oal2f09.sl NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:130477 3',
mRNA sequence.
ACCESSION AA766184
VERSION AA766184.1 GI:2817422
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 318)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1277 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 288.
Location/Qualifiers

FEATURES

```

```

1..318
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1304777"
/clone_lib="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-,
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGGCTCATTTTCTTTTCTTTT-3'
]. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT      71 a      82 c      70 g      95 t
ORIGIN

Query Match      59.5%; Score 298.2; DB 9; Length 318;
Best Local Similarity 99.0%; Pred. No. 1.2e-79;
Matches 300; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGTATGAGCTGTGCCAGGCTTGGGCTCCACATAAGCAGTCTATAGATGCTCTT 60
Db 310 GGTATGAGCTGTGCCAGGCTTGGGCTCCACATAAGCAGTCTATAGATGCTCTT 251

Qy 61 AGGACTGTGCTGTCGACAGCCGCGGCGGAGGCTGCCACACGGAAGCAGATGA 120
Db 250 AGGACTGTGCTGTCGACAGCCGCGGCGGAGGCTGCCACACGGAAGCAGATGA 191

Qy 121 ACTAATTTTCATTTCAAGCAGTTTTTAAAGAGTCTTGGAAACAGACGCGGCACCTTTC 180
Db 190 ACTAATTTTCATTTCAAGCAGTTTTTAAAGAGTCTTGGAAACAGCTTCGCGGCACCTTTC 131

Qy 181 CTCTAATCCAGCAAGTAGTTCCTGCACACGAGACAGCAGAGTAACAGGATCAGTG 240
Db 130 CTCTAATCCAGCAAGTAGTTCCTGCACACGAGACAGCAGAGTAACAGGATCAGTG 71

Qy 241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTACGCTGCAATAAGATTGATTG 300
Db 70 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTACGCTGCAATAAGATTGATTG 11

Qy 301 CAA 303
Db 10 CAA 8

RESULT 14
LOCUS AI033108/c
DEFINITION owa98g08.sl Soares fetal liver spleen INFLS_S1 Homo sapiens cDNA
clone IMAGE:1654910 3', mRNA sequence.
ACCESSION AI033108
VERSION AI033108.1 GI:3254061
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 410)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

```

Insert Length: 773 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 368.

FEATURES

Location/Qualifiers
1..410
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1654910"
/clone_lib="Soares_fetal_liver_spleen_INFLS_s1"
/sex="male"
/dev_stages="D20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;
This is a subcloned version of the original Soares fetal
liver spleen INFLS library. 1st strand cDNA was primed
with a Pac I - oligo(dT) primer [5',
AATGGAGAAATTAATTAAGATCTTTTATTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 88 a 110 c 101 g 111 t
ORIGIN
Query Match 59.5%; Score 298.2; DB 9; Length 410;
Best Local Similarity 99.0%; Pred. No. 1.4e-79;
Matches 300; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 GGTATGGAGCTGTGGCGAGGCTTGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 60
DB 305 GGTATGGAGCTGTGGCGAGGCTTGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 246
QY 61 AGGACTGGTGGCTGGCACAGCGCGGCGAGGCTGCCACACGAGCAAGCAGATGA 120
DB 245 AGGACTGGTGGCTGGCACAGCTGGCGGCGAGGCTGCCACACGAGCAAGCAGATGA 186
QY 121 ACTAATTTTCATTTCAAGGCGAGTCTTTTAAAGAGTCTTTGAAACAGACGCGGCACCTTTC 180
DB 185 ACTAATTTTCATTTCAAGGCGAGTCTTTTAAAGAGTCTTTGAAACAGACGCGGCACCTTTC 126
QY 181 CTCTAATCCAGCAAGTGTATCCCTGCACACGAGACAAAGCAGATGAACAGATCAGTG 240
DB 125 CTCTAATCCAGCAAGTGTATCCCTGCACACGAGACAAAGCAGATGAACAGATCAGTG 66
QY 241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCACTGCAATTAAGATTGAGTTTG 300
DB 65 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCACTGCAATTAAGATTGAGTTTG 6
QY 301 CAA 303
DB 5 CAA 3

RESULT 15
AA223338/c

LOCUS AA2233338 374 bp mRNA linear EST 12-MAR-1998
DEFINITION zr05h05.s1 Stratagene NT2 neuronal precursor 937230 Homo sapiens
CDNA clone IMAGE:650649 3', mRNA sequence.
ACCESSION AA223338
VERSION AA223338.1 GI:1843862
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 374)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin,
J., Moore, B., Schellenberg, K., Stepien, M., Tan, F., Theising, B.,
White, Y., Wyllie, T., Waterston, R. and Willson, R.
WashU-NCI human EST Project

TITLE

JOURNAL
COMMENT

Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
INSE Length: 3011 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham.

FEATURES

Location/Qualifiers
1..374
/organism="Homo sapiens"
/db_xref="GDB:5277238"
/db_xref="taxon:9606"
/clone="IMAGE:650649"
/clone_lib="Stratagene NT2 neuronal precursor 937230"
/tissue_type="neuroepithelial cells"
/dev_stages="Ntera-2 neuroepithelial cells"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: brain; Vector: pBluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. Uninduced, exponentially growing neuroepithelial
cells (Ntera-2/cl.D1). Average insert size: 1.0 kb.
Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGGAG
3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3' "

BASE COUNT 78 a 98 c 85 g 112 t 1 others
ORIGIN

Query Match 59.3%; Score 297.2; DB 9; Length 374;
Best Local Similarity 98.7%; Pred. No. 2.6e-79;
Matches 299; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GGTATGGAGCTGTGGCGAGGCTTGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 60
DB 312 GGTATGGAGCTGTGGCGAGGCTTGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 253
QY 61 AGGACTGGTGGCTGGCACAGCGCGGCGAGGCTGCCACACGAGCAAGCAGATGA 120
DB 252 AGGACTGGTGGCTGGCACAGCTGGCGGCGAGGCTGCCACACGAGCAAGCAGATGA 193
QY 121 ACTAATTTTCATTTCAAGGCGAGTCTTTTAAAGAGTCTTTGAAACAGACGCGGCACCTTTC 180
DB 192 ACTAATTTTCATTTCAAGGCGAGTCTTTTAAAGAGTCTTTGAAACAGACGCGGCACCTTTC 133
QY 181 CTCTAATCCAGCAAGTGTATCCCTGCACACGAGACAAAGCAGATGAACAGATCAGTG 240
DB 132 CTCTAATCCAGCAAGTGTATCCCTGCACACGAGACAAAGCAGATGAACAGATCAGTG 73
QY 241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCACTGCAATTAAGATTGAGTTTG 300
DB 72 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCACTGCAATTAAGATTGAGTTTG 13
QY 301 CAA 303
DB 12 CAA 10

Search completed: May 17, 2003, 18:18:22
Job time : 689.842 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 17:27:39 ; Search time 63.5169 Seconds
(without alignments)
10174.595 Million cell updates/sec

Title: US-09-434-382-28_COPY_26164_26664
Perfect score: 501
Sequence: 1 ggtatgagctgtgccagg.....ctcccaagttttttgaca 501

Scoring table:
Gapop 10.0 , Gapext 1.0

Searched: 810007 seqs, 644969091 residues

Total number of hits satisfying chosen parameters: 1620014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	501	100.0	26664	9	US-09-988-626-28
2	501	100.0	26664	9	US-09-988-687-28
3	303	60.5	655	9	US-09-988-626-27
4	303	60.5	655	9	US-09-988-687-27
5	303	60.5	2958	9	US-09-988-626-3
6	303	60.5	2958	9	US-09-988-687-3
7	296.6	59.2	2908	9	US-09-988-626-223
8	296.6	59.2	2908	9	US-09-988-687-223
9	254.6	50.8	2892	9	US-09-988-626-225
10	254.6	50.8	2892	9	US-09-988-687-225
11	33.6	6.7	480	10	US-09-864-761-15427
12	32.2	6.4	513509	9	US-09-754-853A-4
13	32	6.4	1089	10	US-09-974-300-2320
14	32	6.4	3779	10	US-09-880-107-2145
15	31.4	6.3	2049	10	US-09-887-576-655
16	31.2	6.2	1101	9	US-10-102-806-206
17	31.2	6.2	2327	9	US-09-809-391-107
18	30.8	6.2	529	10	US-09-917-800A-1269
19	30.6	6.1	350	10	US-09-960-352-8591

c 20	30.6	6.1	367	10	US-09-960-352-5310	Sequence 5310, Ap
c 21	30.6	6.1	406	10	US-09-960-352-9329	Sequence 9329, Ap
c 22	30.6	6.1	421	9	US-10-091-504-415	Sequence 415, App
c 23	30.6	6.1	421	10	US-09-764-869-415	Sequence 415, App
c 24	30.6	6.1	463	10	US-09-960-352-1733	Sequence 1733, Ap
c 25	30.6	6.1	38844	12	US-10-060-333-3	Sequence 3, Appli
c 26	30.2	6.0	294	10	US-09-294-093B-2539	Sequence 2539, Ap
c 27	30.2	6.0	728	9	US-10-079-854-54	Sequence 54, Appl
c 28	30.2	6.0	728	10	US-09-764-878-54	Sequence 114, App
c 29	30.2	6.0	1595	9	US-10-108-605-114	Sequence 452, App
c 30	30	6.0	293	10	US-09-294-093B-452	Sequence 704, App
c 31	30	6.0	418	9	US-09-918-995-2942	Sequence 2942, Ap
c 32	30	6.0	515	10	US-09-815-343-704	Sequence 602, App
c 33	30	6.0	779	10	US-09-910-943-602	Sequence 952, App
c 34	30	6.0	1901	10	US-09-834-975-952	Sequence 10, Appl
c 35	30	6.0	4355	12	US-10-044-090-10	Sequence 170, App
c 36	30	6.0	8586	10	US-09-960-253-170	Sequence 806, App
c 37	29.8	5.9	576	10	US-09-924-035A-806	Sequence 6, Appli
c 38	29.8	5.9	180557	12	US-10-003-806-6	Sequence 9, Appli
c 39	29.8	5.9	180557	12	US-10-003-806-9	Sequence 6473, Ap
c 40	29.6	5.9	392	9	US-09-764-891-6473	Sequence 20, Appl
c 41	29.6	5.9	392	9	US-09-764-891-6473	Sequence 72, Appl
c 42	29.6	5.9	544	10	US-09-815-242-20	Sequence 7967, Ap
c 43	29.4	5.9	555	9	US-10-123-155-72	Sequence 2000, Ap
c 44	29.4	5.9	564	10	US-09-864-761-7967	
c 45	29.4	5.9	577	10	US-09-867-701-2000	

ALIGNMENTS

RESULT 1

US-09-988-626-28
; Sequence 28, Application US/09988626
; Publication No. US20030044959A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,626
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 26664
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (910)..(13104)
; OTHER INFORMATION: exon 1: 910-1154; exon 2: 1736-1786; exon 3:
; OTHER INFORMATION: 1925-1995; exon 4: 3025-3089; exon 5: 4361-4418;
; OTHER INFORMATION: exon 6: 5583-5650; exon 7: 7075-7194; exon 8:
; OTHER INFORMATION: 8186-8244; exon 9: 12878-12936; exon 10:
; OTHER INFORMATION: 13032-13104;
; NAME/KEY: misc feature
; LOCATION: (13756)..(22917)
; OTHER INFORMATION: exon 11: 13756-13868; exon 12: 15283-15378; exon 13:
; OTHER INFORMATION: 13: 16278-16416; exon 14: 16498-16583; exon 15:
; OTHER INFORMATION: 18583-18701; exon 16: 20349-20445; exon 17:
; OTHER INFORMATION: 22172-22310; exon 18: 22879-22917
; NAME/KEY: misc feature

```
; LOCATION: (23045)..(26452)
; OTHER INFORMATION: exon 19: 23045-23154; exon 20: 23795-23895; exon
; OTHER INFORMATION: 21: 23973-24093; exon 22: 24354-24432; exon 23:
; OTHER INFORMATION: 25026-25170; exon 24: 25812-26036; polyadenylation
; OTHER INFORMATION: signal: 26447-26452
; NAME/KEY: variation
; LOCATION: (826)..(23879)
; OTHER INFORMATION: s at positions 826 and 23180 is G or C; Y at
; OTHER INFORMATION: positions 1914, 5568, 7165, 16431, 1857 and 20486
; OTHER INFORMATION: is C or T; n at position 13128 is t or tgat; r at
; OTHER INFORMATION: positions 22211 and 23879 is A or G.
US-09-988-626-28

Query Match      100.0%; Score 501; DB 9; Length 26664;
Best Local Similarity 100.0%; Pred. No. 1.7e-160;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 60
Db      26164 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 26223

Qy      61 AGGACTGTGCTGCGCAGACGCGCGGCGAGGCTGCCACCGAAGCAAGCAGATGA 120
Db      26224 AGGACTGTGCTGCGCAGACGCGCGGCGAGGCTGCCACCGAAGCAAGCAGATGA 26283

Qy      121 ACTAATTTCAATTTCAAGCAGTTTTTAAAGAAAGTCTTGAACACAGACGCGGCACTTTTC 180
Db      26284 ACTAATTTCAATTTCAAGCAGTTTTTAAAGAAAGTCTTGAACACAGACGCGGCACTTTTC 26343

Qy      181 CTCTAATCCAGCAAAAGTGATTCCTGTCACACAGAGCAAGCAGAGTAACAGGATCAGTG 240
Db      26344 CTCTAATCCAGCAAAAGTGATTCCTGTCACACAGAGCAAGCAGAGTAACAGGATCAGTG 26403

Qy      241 GGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTACAGTGTCAATAAAGATTGAGTTTG 300
Db      26404 GGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTACAGTGTCAATAAAGATTGAGTTTG 26463

Qy      301 CAATTTGAGTCTTTTGGCTTCTCTGCTGTAGTGTGCGAGCTGCCGTGACCGGGTGGCTTGG 420
Db      26464 CAATTTGAGTCTTTTGGCTTCTCTGCTGTAGTGTGCGAGCTGCCGTGACCGGGTGGCTTGG 26523

Qy      361 ACCTTGGAGAGGCTCTCTGTGCTGTGTAGTGTGCGAGCTGCTGTGACCGGGTGGCTTGG 480
Db      26584 AAGAAAGTCAAGTCCCGTGTAGTGTGAGCACTCTGGAACCTGTGCTCTCAGAGAGCCACCCTT 26643

Qy      481 ATTCGCCAAGTCTTTTGGACA 501
Db      26644 ATTCGCCAAGTCTTTTGGACA 26664

RESULT 2
US-09-988-687-28
; Sequence 28, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigan, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,687
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
```

```
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 28
; LENGTH: 26664
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (910)..(13104)
; OTHER INFORMATION: exon 1: 910-1154; exon 2: 1736-1786; exon 3:
; OTHER INFORMATION: 1925-1995; exon 4: 3025-3089; exon 5: 4361-4418;
; OTHER INFORMATION: exon 6: 5582-5650; exon 7: 7075-7194; exon 8:
; OTHER INFORMATION: 8186-8244; exon 9: 12878-12936; exon 10:
; OTHER INFORMATION: 13032-13104;
; NAME/KEY: misc feature
; LOCATION: (13756)..(22917)
; OTHER INFORMATION: exon 11: 13756-13868; exon 12: 15283-15378; exon
; OTHER INFORMATION: 13: 16278-16416; exon 14: 16498-16583; exon 15:
; OTHER INFORMATION: 18583-18701; exon 16: 20349-20445; exon 17:
; OTHER INFORMATION: 22172-22310; exon 18: 22879-22917
; NAME/KEY: misc feature
; LOCATION: (23045)..(26452)
; OTHER INFORMATION: exon 19: 23045-23154; exon 20: 23795-23895; exon
; OTHER INFORMATION: 21: 23973-24093; exon 22: 24354-24432; exon 23:
; OTHER INFORMATION: 25026-25170; exon 24: 25812-26036; polyadenylation
; OTHER INFORMATION: signal: 26447-26452
; NAME/KEY: variation
; LOCATION: (826)..(23879)
; OTHER INFORMATION: s at positions 826 and 23180 is G or C; Y at
; OTHER INFORMATION: positions 1914, 5568, 7165, 16431, 1857 and 20486
; OTHER INFORMATION: is C or T; n at position 13128 is t or tgat; r at
; OTHER INFORMATION: positions 22211 and 23879 is A or G.
US-09-988-687-28
```

```
Query Match      100.0%; Score 501; DB 9; Length 26664;
Best Local Similarity 100.0%; Pred. No. 1.7e-160;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 60
Db      26164 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 26223

Qy      61 AGGACTGTGCTGCGCAGACGCGCGGCGAGGCTGCCACCGAAGCAAGCAGATGA 120
Db      26224 AGGACTGTGCTGCGCAGACGCGCGGCGAGGCTGCCACCGAAGCAAGCAGATGA 26283

Qy      121 ACTAATTTCAATTTCAAGCAGTTTTTAAAGAAAGTCTTGAACACAGACGCGGCACTTTTC 180
Db      26284 ACTAATTTCAATTTCAAGCAGTTTTTAAAGAAAGTCTTGAACACAGACGCGGCACTTTTC 26343

Qy      181 CTCTAATCCAGCAAAAGTGATTCCTGTCACACAGAGCAAGCAGAGTAACAGGATCAGTG 240
Db      26344 CTCTAATCCAGCAAAAGTGATTCCTGTCACACAGAGCAAGCAGAGTAACAGGATCAGTG 26403

Qy      241 GGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTACAGTGTCAATAAAGATTGAGTTTG 300
Db      26404 GGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTACAGTGTCAATAAAGATTGAGTTTG 26463

Qy      301 CAATTTGAGTCTTTTGGCTTCTCTGCTGTAGTGTGCGAGCTGCCGTGACCGGGTGGCTTGG 360
Db      26464 CAATTTGAGTCTTTTGGCTTCTCTGCTGTAGTGTGCGAGCTGCCGTGACCGGGTGGCTTGG 26523

Qy      361 ACCTTGGAGAGGCTCTCTGTGCTGTGTAGTGTGCGAGCTGCTGTGACCGGGTGGCTTGG 420
Db      26524 ACCTTGGAGAGGCTCTCTGTGCTGTGTAGTGTGCGAGCTGCTGTGACCGGGTGGCTTGG 26583

Qy      421 AAGAAAGTCAAGTCCCGTGTAGTGTGAGCACTCTGGAACCTGTGCTCTCAGAGAGCCACCCTT 480
Db      26584 AAGAAAGTCAAGTCCCGTGTAGTGTGAGCACTCTGGAACCTGTGCTCTCAGAGAGCCACCCTT 26643

Qy      481 ATTCGCCAAGTCTTTTGGACA 501
```


Db 26644 ATTCGCCAAGTCTTTTGACA 26664
|||||

RESULT 3

US-09-988-626-27
; Sequence 27, Application US/09988626
; Publication No. US20030044959A1

GENERAL INFORMATION:

; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Simard, David H.F.
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,626
; CURRENT FILING DATE: 2001-11-20
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 655
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: (1)..(228)
; OTHER INFORMATION: exon 24
; NAME/KEY: polyA_signal
; LOCATION: (636)..(641)
US-09-988-626-27

Query Match 60.5%; Score 303; DB 9; Length 655;
Best Local Similarity 100.0%; Pred. No. 1.8e-93;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTATGGAGCTGTGCGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCCTCTT 60
Db 353 GGTATGGAGCTGTGCGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCCTCTT 412
Qy 61 AGGACTGGTGTGCGACAGCGCGCGGCGGAGGCTGCCACACGGAAGCAAGCAGATGA 120
Db 413 AGGACTGGTGTGCGACAGCGCGCGGCGGAGGCTGCCACACGGAAGCAAGCAGATGA 472
Qy 121 ACTAATTTTCATTTCAAGGAGCTTTTAAAGAGTCTTTGGAACAGACGCGGCACCTTTTC 180
Db 473 ACTAATTTTCATTTCAAGGAGCTTTTAAAGAGTCTTTGGAACAGACGCGGCACCTTTTC 532
Qy 181 CTCCTAATCCAGCAAAAGTATTCCTCCACACGAGAGAGTAAACAGATCAAGATCAGTG 240
Db 533 CTCCTAATCCAGCAAAAGTATTCCTCCACACGAGAGAGTAAACAGATCAAGATCAGTG 592
Qy 241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAAGATTGAGTTTG 300
Db 593 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAAGATTGAGTTTG 652
Qy 301 CAA 303
Db 653 CAA 655

RESULT 4

US-09-988-687-27
; Sequence 27, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:

; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,687
; CURRENT FILING DATE: 2001-11-20
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 655
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: (1)..(228)
; OTHER INFORMATION: exon 24
; NAME/KEY: polyA_signal
; LOCATION: (636)..(641)
US-09-988-687-27

Query Match 60.5%; Score 303; DB 9; Length 655;
Best Local Similarity 100.0%; Pred. No. 1.8e-93;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTATGGAGCTGTGCGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCCTCTT 60
Db 353 GGTATGGAGCTGTGCGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCCTCTT 412
Qy 61 AGGACTGGTGTGCGACAGCGCGCGGCGGAGGCTGCCACACGGAAGCAAGCAGATGA 120
Db 413 AGGACTGGTGTGCGACAGCGCGCGGCGGAGGCTGCCACACGGAAGCAAGCAGATGA 472
Qy 121 ACTAATTTTCATTTCAAGGAGCTTTTAAAGAGTCTTTGGAACAGACGCGGCACCTTTTC 180
Db 473 ACTAATTTTCATTTCAAGGAGCTTTTAAAGAGTCTTTGGAACAGACGCGGCACCTTTTC 532
Qy 181 CTCCTAATCCAGCAAAAGTATTCCTCCACACGAGAGAGTAAACAGATCAAGATCAGTG 240
Db 533 CTCCTAATCCAGCAAAAGTATTCCTCCACACGAGAGAGTAAACAGATCAAGATCAGTG 592
Qy 241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAAGATTGAGTTTG 300
Db 593 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAAGATTGAGTTTG 652
Qy 301 CAA 303
Db 653 CAA 655

RESULT 5

US-09-988-626-3
; Sequence 3, Application US/09988626
; Publication No. US20030044959A1

GENERAL INFORMATION:

; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,626

QY 1 GGTATGGAGCTGTGCGAGGCTTGGCTCCACATTAAGCACTAGTCTATAGATGCCTCTT 60
Db 2606 GGTATGGAGCTGTGCGAAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCCTCTT 2665

QY 61 AGGACTGGTGCCTGGGCACAGCCGGCCAGAGGCTGCCACACGGAAGCAAGCAGATGA 120
Db 2666 AGGACTGGTGCCTGGGCACAGCCGGGCAGAGGCTGCCACACGGAAGCAAGCAGATGA 2725

QY 121 ACTAATTTTCATTTCAAGGAGCTTTTAAAGAGTCTTTGAAACAGACGCGGCACCTTTC 180
Db 2726 ACTAATTTTCATTTCAAGGAGCTTTTAAAGAGGCTTTGAAACAGACGCGGCACCTTTC 2785

QY 181 CTCTAATCCAGCAAGTGAATCCCTGCACACAGACAGCAAGCAGATTAAGATCAGTG 240
Db 2786 CTCTAATCCAGCAAGTGAATCCCTGCACACAGACAGCAAGCAGATTAAGATCAGTG 2845

QY 241 GGTCTAAGTGTCCGAGACTTAAAGAAATAGTATTTTCAAGTCAATTAAGATTTGAGTTTG 300
Db 2846 GGTCTAAGTGTCCGAGACTTAAAGAAATAGTATTTTCAAGTCAATTAAGATTTGAGTTTG 2905

QY 301 CAA 303
Db 2906 CAA 2908

RESULT 8

US-09-988-687-223
; Sequence 223, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,687
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 223
; LENGTH: 2908
; TYPE: DNA
; ORGANISM: Pan troglodytes
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
US-09-988-687-223

Query Match 59.2%; Score 296.6; DB 9; Length 2908;
Best Local Similarity 98.7%; Pred No. 6.7e-91;
Matches 299; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGTATGGAGCTGTGCGAGGCTTGGCTCCACATAAGCACTAGTCTATAGATGCCTCTT 60
Db 2606 GGTATGGAGCTGTGCGAAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCCTCTT 2665

QY 61 AGGACTGGTGCCTGGGCACAGCCGGCCAGAGGCTGCCACACGGAAGCAAGCAGATGA 120
Db 2666 AGGACTGGTGCCTGGGCACAGCCGGGCAGAGGCTGCCACACGGAAGCAAGCAGATGA 2725

QY 121 ACTAATTTTCATTTCAAGGAGCTTTTAAAGAGTCTTTGAAACAGACGCGGCACCTTTC 180
Db 2726 ACTAATTTTCATTTCAAGGAGCTTTTAAAGAGGCTTTGAAACAGACGCGGCACCTTTC 2785

QY 181 CTCTAATCCAGCAAGTGAATCCCTGCACACAGACAGCAAGCAGATTAAGATCAGTG 240
Db 2786 CTCTAATCCAGCAAGTGAATCCCTGCACACAGACAGCAAGCAGATTAAGATCAGTG 2845

QY 241 GGTCTAAGTGTCCGAGACTTAAAGAAATAGTATTTTCAAGTCAATTAAGATTTGAGTTTG 300
Db 2846 GGTCTAAGTGTCCGAGACTTAAAGAAATAGTATTTTCAAGTCAATTAAGATTTGAGTTTG 2905

QY 301 CAA 303
Db 2906 CAA 2908

RESULT 9
US-09-988-626-225
; Sequence 225, Application US/09988626
; Publication No. US20030044959A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,626
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 225
; LENGTH: 2892
; TYPE: DNA
; ORGANISM: Gorilla gorilla
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
US-09-988-626-225

Query Match 50.8%; Score 254.6; DB 9; Length 2892;
Best Local Similarity 93.4%; Pred No. 1.7e-76;
Matches 283; Conservative 0; Mismatches 4; Indels 16; Gaps 1;

QY 1 GGTATGGAGCTGTGCGAGGCTTGGCTCCACATAAGCACTAGTCTATAGATGCCTCTT 60
Db 2606 GGTATGGAGCTGTGCGAGGCTTAGGCTCCACATAAGCACTAGTCTATA----- 2655

QY 61 AGGACTGGTGCCTGGGCACAGCCGGCCAGAGGCTGCCACACGGAAGCAAGCAGATGA 120
Db 2656 -----GGTGCCTGGCAGACGCGCGGAGGCTGCCACACGGAAGCAAGCAGATGA 2709

QY 121 ACTAATTTTCATTTCAAGCAGTCTTTTAAAGAGTCTTTGAAACAGACGCGGCACCTTTC 180
Db 2710 ACTAATTTTCATTTCAAGCAGTCTTTTAAAGAGTCTTTGAAACAGACGCGGCACCTTTC 2769

QY 181 CTCTAATCCAGCAAGTGAATCCCTGCACACAGACAGCAAGCAGATTAAGATCAGTG 240
Db 2770 CTCTAATCCAGCAAGTGAATCCCTGCACACAGACAGCAAGCAGATTAAGATCAGTG 2829

QY 241 GGTCTAAGTGTCCGAGACTTAAAGAAATAGTATTTTCAAGTCAATTAAGATTTGAGTTTG 300
Db 2830 GGTCTAAGTGTCCGAGACTTAAAGAAATAGTATTTTCAAGTCAATTAAGATTTGAGTTTG 2889

QY 301 CAA 303
Db 2890 CAA 2892

```
RESULT 10
US-09-988-687-225
; Sequence 225, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: 17p-Linked prostate Cancer Susceptibility
; TITLE OF INVENTION: Chromosome 17 and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,687
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 225
; LENGTH: 2892
; TYPE: DNA
; ORGANISM: Gorilla gorilla
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
US-09-988-687-225

Query Match      50.8%; Score 254.6; DB 9; Length 2892;
Best Local Similarity 93.4%; Pred. No. 1.7e-76;
Matches 283; Conservative 0; Mismatches 4; Indels 16; Gaps 1;

Qy 1 GGTATGGAGCTGTGCGGAGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 60
Db 2606 GGTATGGAGCTGTGCGGAGCTTAGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 60

Qy 61 AGGACTGTGCTGGCAGACGCGGGCGGAGGCTGCCACAGCAAGCAAGCAGATGA 120
Db 2656 -----GGTGCTGTGCAGACGCGGGCAGGAGCTGCCACAGCAAGCAAGCAGATGA 2709

Qy 121 ACTAATTTCAATTTAAGGAGCTTTTAAAGAGCTTTTGAAGACAGACGCGGCACCTTTC 180
Db 2710 ACTAATTTCAATTTAAGGAGCTTTTAAAGAGCTTTTGAAGACAGACGCGGCACCTTTC 2769

Qy 181 CTCTAATCAGCAAGATGATTCCTTGCACACAGACAGACAGATGAACAGGATCAGTG 240
Db 2770 CTCTAATCAGCAAGATGATTCCTTGCACACAGACAGACAGATGAACAGGATCAGTG 2829

Qy 241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGTGTCAATAAGATTGAGTTTG 300
Db 2830 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGTGTCAATAAGATTGAGTTTG 2889

Qy 301 CAA 303
Db 2890 CAA 2892

RESULT 11
US-09-864-761-15427/c
; Sequence 15427, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
```

```
FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 15427
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009230.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.99
US-09-864-761-15427

Query Match      6.7%; Score 33.6; DB 10; Length 480;
Best Local Similarity 47.2%; Pred. No. 0.44;
Matches 102; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

Qy 176 CTTTCCTCTAATCCAGCAAGTAGTTCCTTCGCACACAGACAGACAGATTAACAGAT 235
Db 344 CTTGAAGCTATTTCGAAATGGTGTAGCTTTTAAAGAAAAAGCAAGTGAACACAGTAA 285

Qy 236 CAGTGGGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGCTCAATAAAGATTGA 295
Db 284 ACAGAGACAAAAAATTTTGGAAACATCAACATGGTACAAATGACCACTTTTAACAGATGA 225

Qy 296 GTTTGCAATTTGTAGTGTCTTTTGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 355
Db 224 AACAAACTTATAGCTTTTCATCTTACAATTACTGTTTCTAGAGAGCTAGATTACTCA 165

Qy 356 GCACACCTTGGAGAAAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 391
Db 164 TCAGGCTTTTGGGGAATTTTCTTTTATTCCTTTTGT 129
```

```

RESULT 12
US-09-754-853A-4
; Sequence 4, Application US/09754853A
; Publication No. US20030005491A1
; GENERAL INFORMATION:
; APPLICANT: Hauge, Brian M.
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Parsons, Jeremy D.
; APPLICANT: Wang, Ming Li
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-10(15810)B
; CURRENT APPLICATION NUMBER: US/09/754,853A
; CURRENT FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 60/174,880
; PRIOR FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 1119
; SEQ ID NO 4
; LENGTH: 513509
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (111805)..(113968), (114684)..(115204)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(513509)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: 318013_region_A3
US-09-754-853A-4

Query Match          6.4%; Score 32.2; DB 9; Length 513509;
Best Local Similarity 66.7%; Pred. No. 69;
Matches 46; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 3 TATGGAGTGTCCGAGGCTGGCTCCACATAGCACTAGTCTATATAGCTCTTAG 62
DB 503895 TTGTGAAGTTGTGCAAGAGTGTGTCTCACACAAAGGCTTAGTACATTGAGTGTGTTAG 503954

QY 63 GACTGGTGC 71
DB 503955 TACTGTGC 503963

RESULT 13
US-09-974-300-2320/c
; Sequence 2320, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2320
; LENGTH: 1089
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-2320

Query Match          6.4%; Score 32; DB 10; Length 1089;
Best Local Similarity 56.7%; Pred. No. 2.5;
Matches 59; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 270 AGTATTTCAGTCGCAATAAAGATTGATTGCAATTGTGAGTCTTTTGTCTCTCTCTGC 329

```

```

DB 527 ATTCTCCCATCTATATAACACACTTGTGTCTAAATTTTCAGCTTTTGTCTTCTCCGA 468
QY 330 TGCTGTCTACAGAGCAGGCTCTCTGTGCAACACACCTTGGAGAAGG 373
DB 467 CGCTCTTCTCAGCAGAACCTTTTCTTCTCGATGTAAAAAG 424

RESULT 14
US-09-880-107-2145
; Sequence 2145, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2145
; LENGTH: 3779
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 L07077
US-09-880-107-2145

Query Match          6.4%; Score 32; DB 10; Length 3779;
Best Local Similarity 54.2%; Pred. No. 5;
Matches 65; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 132 TTCAGGCGAGTCTTTTAAAGAGTCTTGGAAACAGACGGCGGACCTTCTCTAATCCAG 191
DB 1422 TTGTAGGCACTGTTTGGATTGTGGGAATCGAATGTTGNATCTTACTACATCAGG 1481

QY 192 CAAAGTGATTCCTGCACACAGAGACAGAGATACAGGATCAGTGGGTCTAAGTGT 251
DB 1482 CATATTTCTTGTAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1541

RESULT 15
US-09-887-576-655/c
; Sequence 655, Application US/09887576
; Patent No. US20020144047A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.
; APPLICANT: Brown, D.
; APPLICANT: Chang, H.
; APPLICANT: Zhu, T.
; APPLICANT: Han, B.
; APPLICANT: Wang, X.
; APPLICANT: Cooper, Bret
; TITLE OF INVENTION: Promoters for regulation of plant expression
; FILE REFERENCE: 1360.001US1
; CURRENT APPLICATION NUMBER: US/09/887,576
; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/213,848
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214,087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/258,692
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 655

```

```
; LENGTH: 2049
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-887-576-655

Query Match      6.3%; Score 31.4; DB 10; Length 2049;
Best Local Similarity 56.2%; Pred. No. 5.7;
Matches 59; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy  274 TTTCAGCTGCAATAAAGATTGAGTTTGCAATTGTGAGTTCCTTTTGGCTTCCTCCTCCTGCT 333
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db  116 TTTATGTCTGTAAGGACCATTTGGCTTTGAAAAAATCAGATCTCTGAATGCAGTTGGTGCA 57

Qy  334 GCTACAGACGAGGCTCTGCTGTGCACCACTTGGAGAGGCTCTC 378
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db   56 GCGACAGCACAGAGCTTATGACCACAAACCAGAAAGATTGCTCGC 12
```

Search completed: May 18, 2003, 00:39:58
Job time : 384.517 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 15:56:54 ; Search time 28.1908 Seconds
(without alignments)
5450.193 Million cell updates/sec

Title: US-09-434-382-28_COPY_26164_26664

Perfect score: 501

Sequence: 1 ggtatggagctgtgcgagg.....ttcgcaagttttttgaca 501

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata/2/ina/5A COMB.seq:*
- 2: /cgn2_6/prodata/2/ina/5B COMB.seq:*
- 3: /cgn2_6/prodata/2/ina/6A COMB.seq:*
- 4: /cgn2_6/prodata/2/ina/6B COMB.seq:*
- 5: /cgn2_6/prodata/2/ina/PTUS COMB.seq:*
- 6: /cgn2_6/prodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	501	100.0	26564	4	US-09-564-805-28
2	303	60.5	655	4	US-09-564-805-27
3	303	60.5	2558	4	US-09-564-805-3
4	296.6	59.2	2908	4	US-09-564-805-223
5	254.6	50.8	2892	4	US-09-564-805-225
6	33.2	6.6	6042	1	US-08-261-822A-1
7	33.2	6.6	6042	5	PCT-US95-07744A-1
8	33.2	6.6	6172	2	US-08-819-288-1
9	33.2	6.6	6172	4	US-09-400-348-1
10	31.2	6.2	2327	4	US-09-149-476-107
11	31.2	6.2	8916	4	US-09-579-181-11
12	31.2	6.2	9354	4	US-09-579-181-10
13	30.6	6.1	787	4	US-08-943-731-200
14	30.6	6.1	20084	4	US-08-943-731-5
15	30.6	6.1	38844	4	US-09-734-675-3
16	30.2	6.0	35828	4	US-09-449-218D-17
17	30	6.0	1834	1	US-08-592-126-90
18	29.8	5.9	7218	1	US-08-232-463-14
19	29.6	5.9	630	4	US-08-180-371-17
20	29.6	5.9	1677	2	US-08-684-101-1
21	29.6	5.9	1677	4	US-09-205-814-1
22	29.4	5.9	1342	3	US-09-961-083-181
23	29.4	5.9	3048	1	US-08-188-228-47
24	29.4	5.9	3048	1	US-08-332-643-41
25	29.4	5.9	3048	1	US-08-332-638-47
26	29.4	5.9	4453	2	US-08-843-530B-17
27	29.4	5.9	6867	4	US-08-961-527-192

c 28	29	5.8	506	1	US-08-469-802B-7	Sequence 7, Appli
c 29	29	5.8	506	2	US-08-267-803B-7	Sequence 7, Appli
c 30	29	5.8	49795	4	US-09-453-702B-60	Sequence 60, Appli
c 31	28.8	5.7	847	4	US-09-142-565-5	Sequence 5, Appli
c 32	28.8	5.7	3186	1	US-08-761-258-8	Sequence 8, Appli
c 33	28.8	5.7	3186	2	US-08-977-306-8	Sequence 8, Appli
c 34	28.8	5.7	5053	2	US-08-685-576-2	Sequence 2, Appli
c 35	28.6	5.7	3395	4	US-09-103-478-3	Sequence 3, Appli
c 36	28.6	5.7	3395	4	US-09-193-931C-3	Sequence 4, Appli
c 37	28.6	5.7	7560	4	US-09-103-478-4	Sequence 4, Appli
c 38	28.6	5.7	7560	4	US-09-193-931C-4	Sequence 4, Appli
c 39	28.6	5.7	11531	1	US-08-068-945A-1	Sequence 1, Appli
c 40	28.6	5.7	11531	1	US-08-442-806-1	Sequence 1, Appli
c 41	28.4	5.7	877	2	US-08-394-152A-44	Sequence 44, Appli
c 42	28.2	5.6	391	4	US-09-641-638-419	Sequence 419, App
c 43	28.2	5.6	2718	4	US-09-651-656-14	Sequence 14, Appl
c 44	28.2	5.6	2718	4	US-09-650-855-14	Sequence 14, Appl
c 45	28	5.6	603	4	US-09-134-001C-1226	Sequence 1226, Ap

ALIGNMENTS

RESULT 1

US-09-564-805-28

; Sequence 28, Application US/09564805

; Patent No. 6333403

; GENERAL INFORMATION:

; APPLICANT: Tavtigan, Sean V.

; APPLICANT: Teng, David H.F.

; APPLICANT: Simard, Jacques

; APPLICANT: Rommens, Johanna M.

; APPLICANT: Myriad Genetics, Inc.

; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility

; FILE OF INVENTION: Gene and a Paralog and Orthologous Genes

; FILE REFERENCE: 2318-258

; CURRENT APPLICATION NUMBER: US/09/564,805

; CURRENT FILING DATE: 2000-05-05

; PRIOR APPLICATION NUMBER: US 60/107,468

; PRIOR FILING DATE: 1998-11-06

; PRIOR APPLICATION NUMBER: 09/434,382

; PRIOR FILING DATE: 1999-11-05

; NUMBER OF SEQ ID NOS: 240

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 28

; LENGTH: 26564

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (910)..(13104)

; OTHER INFORMATION: exon 1: 910-1154; exon 2: 1736-1786; exon 3:

; OTHER INFORMATION: 1925-1995; exon 4: 3025-3089; exon 5: 4361-4418;

; OTHER INFORMATION: exon 6: 5582-5850; exon 7: 7075-7194; exon 8:

; OTHER INFORMATION: 8186-8244; exon 9: 12878-12936; exon 10:

; NAME/KEY: misc feature

; LOCATION: (13756)..(22917)

; OTHER INFORMATION: exon 11: 13756-13868; exon 12: 15283-15378; exon

; OTHER INFORMATION: 13: 16278-16416; exon 14: 16498-16583; exon 15:

; OTHER INFORMATION: 18583-18701; exon 16: 20349-20445; exon 17:

; OTHER INFORMATION: 22172-22310; exon 18: 22879-22917

; NAME/KEY: misc feature

; LOCATION: (23045)..(26452)

; OTHER INFORMATION: exon 19: 23045-23154; exon 20: 23795-23895; exon

; OTHER INFORMATION: 21: 23973-24093; exon 22: 24354-24432; exon 23:

; OTHER INFORMATION: 25026-25170; exon 24: 25812-26036; polyadenylation

; OTHER INFORMATION: signal: 26447-26452

; NAME/KEY: variation

; LOCATION: (826)..(23879)

; OTHER INFORMATION: s at positions 826 and 23180 is G or C; y at

; OTHER INFORMATION: positions 1314, 5568, 7165, 16431, 1857 and 20486

; OTHER INFORMATION: is C or T; n at position 13128 is t or tgat; r at

; OTHER INFORMATION: positions 2221 and 23879 is A or G.

US-09-564-805-28

```
Query Match      100.0%; Score 501; DB 4; Length 26664;
Best Local Similarity 100.0%; Pred. No. 6.1e-160;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTATGGAGCTGTGCGGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 60
Db 26164 GGTATGGAGCTGTGCGGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 26223

Qy 61 AGGACTGGTGTGCTGGCAGACAGCGCGGCGGAGGCTGCCACACGGAAGCAAGCAGATGA 120
Db 26224 AGGACTGGTGTGCTGGCAGACAGCGCGGCGGAGGCTGCCACACGGAAGCAAGCAGATGA 26283

Qy 121 ACTAATTTCAATTTCAAGGAGCTTTTAAAGAAAGTCTTGGAAAACAGACGGCGCACCTTTC 180
Db 26284 ACTAATTTCAATTTCAAGGAGCTTTTAAAGAAAGTCTTGGAAAACAGACGGCGCACCTTTC 26343

Qy 181 CTCTAATCCAGCAAGTGTATCCCTGCACACAGAGAGCAAGCAGAGTAAACAGATCAGTG 240
Db 26344 CTCTAATCCAGCAAGTGTATCCCTGCACACAGAGAGCAAGCAGAGTAAACAGATCAGTG 26403

Qy 241 GGTCTAAGTGTCCGAGACTTAAACGAAAATAGTATTTTCAAGTCTCAATTAAGATTGAGTTTG 300
Db 26404 GGTCTAAGTGTCCGAGACTTAAACGAAAATAGTATTTTCAAGTCTCAATTAAGATTGAGTTTG 26463

Qy 301 CAATTGTGAGTTCTTTTGTCTCTCTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 360
Db 26464 CAATTGTGAGTTCTTTTGTCTCTCTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 26523

Qy 361 ACCTTGGAGAGGCTCTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 420
Db 26524 ACCTTGGAGAGGCTCTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 26583

Qy 421 AAGAAGTCTAGCTCCCGTCTAGTGTAGCAGCTCTGGAACCTGTCTCTCAGAGAGCCACCTTT 480
Db 26584 AAGAAGTCTAGCTCCCGTCTAGTGTAGCAGCTCTGGAACCTGTCTCTCAGAGAGCCACCTTT 26643

Qy 481 ATTGCGCAAGTCTTTTGACA 501
Db 26644 ATTGCGCAAGTCTTTTGACA 26664
```

RESULT 2

```
US-09-564-805-27
; Sequence 27, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 27
; LENGTH: 655
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(228)
; OTHER INFORMATION: exon 24
; NAME/KEY: polyA_signal
```

; LOCATION: (636)..(641)
US-09-564-805-27

```
Query Match      60.5%; Score 303; DB 4; Length 655;
Best Local Similarity 100.0%; Pred. No. 2.1e-93;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTATGAGCTGTGCGGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 60
Db 353 GGTATGAGCTGTGCGGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 412

Qy 61 AGGACTGGTGTGCTGGCAGACAGCGCGGCGGAGGCTGCCACACGGAAGCAAGCAGATGA 120
Db 413 AGGACTGGTGTGCTGGCAGACAGCGCGGCGGAGGCTGCCACACGGAAGCAAGCAGATGA 472

Qy 121 ACTAATTTCAATTTCAAGGAGCTTTTAAAGAAAGTCTTGGAAAACAGACGGCGCACCTTTC 180
Db 473 ACTAATTTCAATTTCAAGGAGCTTTTAAAGAAAGTCTTGGAAAACAGACGGCGCACCTTTC 532

Qy 181 CTCTAATCCAGCAAGTGTATCCCTGCACACAGAGAGCAAGCAGATTAACAGGATCAGTG 240
Db 533 CTCTAATCCAGCAAGTGTATCCCTGCACACAGAGAGCAAGCAGATTAACAGGATCAGTG 592

Qy 241 GGTCTAAGTGTCCGAGACTTAAACGAAAATAGTATTTTCAAGTCTCAATTAAGATTGAGTTTG 300
Db 593 GGTCTAAGTGTCCGAGACTTAAACGAAAATAGTATTTTCAAGTCTCAATTAAGATTGAGTTTG 652

Qy 301 CAA 303
Db 653 CAA 655
```

RESULT 3

```
US-09-564-805-3
; Sequence 3, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2958
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51)..(2531)
; OTHER INFORMATION: coding sequence as in SEQ ID NO:1
```

US-09-564-805-3

```
Query Match      60.5%; Score 303; DB 4; Length 2958;
Best Local Similarity 100.0%; Pred. No. 5.1e-93;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTATGAGCTGTGCGGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 60
Db 2656 GGTATGAGCTGTGCGGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 2715

Qy 61 AGGACTGGTGTGCTGGCAGACAGCGCGGCGGAGGCTGCCACACGGAAGCAAGCAGATGA 120
Db 2656 AGGACTGGTGTGCTGGCAGACAGCGCGGCGGAGGCTGCCACACGGAAGCAAGCAGATGA 120
```



```
Db 2716 AGGACTGGTCCCTGGCACAGCCGCGGCGCAGGAGGCTGCCACAGGAAGCAAGCAGATGA 2775
QY 121 ACTAATTTTCATTTCAAGGAGTCTTTTAAAGAGTCTTTGAAACAGACGCGGCGCACCTTTC 180
Db 2776 ACTAATTTTCATTTCAAGGAGTCTTTTAAAGAGTCTTTGAAACAGACGCGGCGCACCTTTC 2835
QY 181 CTCCTAATCCAGCAAGTGTTCCTTCGCACACAGACAGCAAGCAGAGATTAACAGGATCAGTG 240
Db 2836 CTCCTAATCCAGCAAGTGTTCCTTCGCACACAGACAGCAAGCAGAGATTAACAGGATCAGTG 2895
QY 241 GGTCTAAGTGTCCGAGACTTAAACGAAATAAGTATTTTTCAGCTGCAATAAAGATTGAGTTTG 300
Db 2896 GGTCTAAGTGTCCGAGACTTAAACGAAATAAGTATTTTTCAGCTGCAATAAAGATTGAGTTTG 2955
QY 301 CAA 303
Db 2956 CAA 2958

RESULT 4
US-09-564-805-223
; Sequence 223, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 223
; LENGTH: 2908
; TYPE: DNA
; ORGANISM: Pan troglodytes
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
US-09-564-805-223

Query Match 59.2%; Score 296.6; DB 4; Length 2908;
Best Local Similarity 98.7%; Pred. No. 7.6e-91;
Matches 299; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGTATGAGCTGTGCCAGGCTTGGGCTCCACATAGCACTAGTCTATAGATGCCTCTT 60
Db 2606 GGTATGAGCTGTGCCAGGCTTGGGCTCCACATAGCACTAGTCTATAGATGCCTCTT 2665
QY 61 AGGACTGGTCCCTGGCACAGCCGCGGCGCAGGAGGCTGCCACAGGAAGCAAGCAGATGA 120
Db 2666 AGGACTGGTCCCTGGCACAGCCGCGGCGCAGGAGGCTGCCACAGGAAGCAAGCAGATGA 2725
QY 121 ACTAATTTTCATTTCAAGGAGTCTTTTAAAGAGTCTTTTAAAGAGTCTTTTAAAGAGTCTTTT 180
Db 2726 ACTAATTTTCATTTCAAGGAGTCTTTTAAAGAGTCTTTTAAAGAGTCTTTTAAAGAGTCTTTT 2785
QY 181 CTCCTAATCCAGCAAGTGTTCCTTCGCACACAGACAGCAAGCAGAGATTAACAGGATCAGTG 240
Db 2786 CTCCTAATCCAGCAAGTGTTCCTTCGCACACAGACAGCAAGCAGAGATTAACAGGATCAGTG 2845
QY 241 GGTCTAAGTGTCCGAGACTTAAACGAAATAAGTATTTTTCAGCTGCAATAAAGATTGAGTTTG 300
Db 2846 GGTCTAAGTGTCCGAGACTTAAACGAAATAAGTATTTTTCAGCTGCAATAAAGATTGAGTTTG 2905
```

```
QY 301 CAA 303
Db 2906 CAA 2908

RESULT 5
US-09-564-805-225
; Sequence 225, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 225
; LENGTH: 2892
; TYPE: DNA
; ORGANISM: Gorilla gorilla
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
US-09-564-805-225

Query Match 50.8%; Score 254.6; DB 4; Length 2892;
Best Local Similarity 93.4%; Pred. No. 1.6e-76;
Matches 283; Conservative 0; Mismatches 4; Indels 16; Gaps 1;

QY 1 GGTATGAGCTGTGCCAGGCTTGGGCTCCACATAGCACTAGTCTATAGATGCCTCTT 60
Db 2606 GGTATGAGCTGTGCCAGGCTTGGGCTCCACATAGCACTAGTCTATAGATGCCTCTT 2655
QY 61 AGGACTGGTCCCTGGCACAGCCGCGGCGCAGGAGGCTGCCACAGGAAGCAAGCAGATGA 120
Db 2656 -----GGTGCCTGGCACAGCCGCGGCGCAGGAGGCTGCCACAGGAAGCAAGCAGATGA 2709
QY 121 ACTAATTTTCATTTCAAGGAGTCTTTTAAAGAGTCTTTTAAAGAGTCTTTTAAAGAGTCTTTT 180
Db 2710 ACTAATTTTCATTTCAAGGAGTCTTTTAAAGAGTCTTTTAAAGAGTCTTTTAAAGAGTCTTTT 2769
QY 181 CTCCTAATCCAGCAAGTGTTCCTTCGCACACAGACAGCAAGCAGAGATTAACAGGATCAGTG 240
Db 2770 CTCCTAATCCAGCAAGTGTTCCTTCGCACACAGACAGCAAGCAGAGATTAACAGGATCAGTG 2829
QY 241 GGTCTAAGTGTCCGAGACTTAAACGAAATAAGTATTTTTCAGCTGCAATAAAGATTGAGTTTG 300
Db 2830 GGTCTAAGTGTCCGAGACTTAAACGAAATAAGTATTTTTCAGCTGCAATAAAGATTGAGTTTG 2889
QY 301 CAA 303
Db 2890 CAA 2892

RESULT 6
US-08-261-822A-1
; Sequence 1, Application US/08261822A
; Patent No. 5650553
; GENERAL INFORMATION:
; APPLICANT: Ecker, Joseph R. et al.
; TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene
; TITLE OF INVENTION: and Pathogens
; NUMBER OF SEQUENCES: 82
```

```
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5650553ris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/261.822A
; FILING DATE: 17-JUN-1994
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Beardell, Lori Y.
; REGISTRATION NUMBER: 34,293
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6042 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
;
; US-08-261-822A-1
;
; Query Match 6.6%; Score 33.2; DB 1; Length 6042;
; Best Local Similarity 51.3%; Pred. No. 0.67;
; Matches 77; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
;
; QY 200 TTCCTGTCACACGACAGACAGGATACAGGATCAGTGGGTCTAAGTGTCCGAGACT 259
;   |||||
; Db 618 TTCCTTGAAGATCTGAATCGGTAGATCATACGGGATCTTTGCAATTTTGTCTTTTCGT 677
;
; QY 260 TAACGAAAATAGTATTTCAGCTGCAATAAGATTGAGTTTGCAAATGTGAGTCTTTTGC 319
;   |||||
; Db 678 CAGCGTTACGATCTTTAGCTTCAGTTAGTTGAAATTTGATTTTTCAGCTTATC 737
;
; QY 320 TTCCTCTGCTGCTGTACAGACAGGGTC 349
;   |||||
; Db 738 TTCCTTTTGTGCTGCTTCATACATAGATC 767
;
;
; RESULT 7
; PCT-US95-07744A-1
; Sequence 1, Application PC/TUS9507744A
; GENERAL INFORMATION:
; APPLICANT: Trustees of The University of Pennsylvania
; TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene
; TITLE OF INVENTION: and Pathogens
; NUMBER OF SEQUENCES: 82
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & Norris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07744A
; FILING DATE: 15-JUNE-1995
;
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5650553ris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07744A
; FILING DATE: 15-JUNE-1995
;
;
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/261.822
; FILING DATE: June 17, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Beardell, Lori Y.
; REGISTRATION NUMBER: 34,293
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6042 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
;
; PCT-US95-07744A-1
;
; Query Match 6.6%; Score 33.2; DB 5; Length 6042;
; Best Local Similarity 51.3%; Pred. No. 0.67;
; Matches 77; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
;
; QY 200 TTCCTGTCACACGACAGACAGGATACAGGATCAGTGGGTCTAAGTGTCCGAGACT 259
;   |||||
; Db 618 TTCCTTGAAGATCTGAATCGGTAGATCATACGGGATCTTTGCAATTTTGTCTTTTCGT 677
;
; QY 260 TAACGAAAATAGTATTTCAGCTGCAATAAGATTGAGTTTGCAAATGTGAGTCTTTTGC 319
;   |||||
; Db 678 CAGCGTTACGATCTTTAGCTTCAGTTAGTTGAAATTTGATTTTTCAGCTTATC 737
;
; QY 320 TTCCTCTGCTGCTGTACAGACAGGGTC 349
;   |||||
; Db 738 TTCCTTTTGTGCTGCTTCATACATAGATC 767
;
;
; RESULT 8
; US-08-819-288-1
; Sequence 1, Application US/08819288
; GENERAL INFORMATION:
; APPLICANT: Ecker, Joseph
; APPLICANT: Alonso, Jose
; TITLE OF INVENTION: PLANT GENES FOR SENSITIVITY TO ETHYLENE
; TITLE OF INVENTION: AND PATHOGENS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5955652ris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/819,288
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Beardell, Lori Y.
; REGISTRATION NUMBER: 34,293
; REFERENCE/DOCKET NUMBER: UPN-2949
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
```

```

, LENGTH: 6172 base pairs
, TYPE: nucleic acid
, STRANDEDNESS: single
, TOPOLOGY: linear
, MOLECULE TYPE: DNA (genom
, HYPOTHETICAL: NO
, ANTI-SENSE: NO
US-08-819-238-1

```

Query Match	6.6%	Score 33.2	DB 2	Length 6172
Best Local Similarity	51.3%	Pred. No. 0.68		
Matches 77	Conservative	0	Mismatches 73	Indels 0
Gaps	0			
Qy	200	TTCCCTGCACACCGAGACAAGCAGAGTAACAGGATCAGTGGGTCTAAGTGTCCGAGACT	259	
Db	750	TTCCCTTGAAGATCTGAATGCGTAGATCATACGGGATCTTTCGATTTTGTGTGCTTTTCGT	809	
Qy	260	TAAACGAAATAGTATTTTCAGCTGCGCAATAAAGATTGAGTTTGCAGTTCTTTTGC	319	
Db	810	CAGCGTTACGATCTTTTAGCTTCAGTTTAGTTGAAATTTGATTTTTTTTCAGCTTATC	869	
Qy	320	TTCTCTCTGCTGCTGCTACAGAGCAGGGTC	349	
Db	870	TTCTTTTTTGTGCTGCTTCATAAGATC	899	

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

Query Match          6.6%; Score 33.2; DB 4; Length 6172;
Best Local Similarity 51.3%; Pred. No. 0.68;
Matches 77; Conservative 0; Mismatches 73; Indels 0; Gaps
0

QY 200 TTCCCTCCACACAGAGACAAGCAGAGTACAGAGATCAGTGGGTCTAAAGTGTCCGAGACT 259
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 750 TTCTCTGAAGATCTGAATGCGTAGATCATACGGGATCTTTTGCATTTTGTGCTTTTCGT 809

QY 260 TAAGCAAAATAGTATTTCAGCTGCCAATAAGAGATTGAGTTTGCAAATTGTGAGTTCTTTTGC 319
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 810 CAGCGTTTACGATTCCTTTTACGCTTCAGTTAGTTGAAATTTGTATTATTTTGTGAGCTTATC 869

QY 320 TTCTCTCTGCTGCTGTACAGACGAGGTC 349
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 870 TTCTTTTGTGTGCTTTCATCTACTAAGATC 899

RESULT 10
US-09-149-476-107
; Sequence 107, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23

```

1	EARLIER APPLICATION NUMBER: 60/047,590
2	EARLIER FILING DATE: 1997-05-23
3	EARLIER APPLICATION NUMBER: 60/047,613
4	EARLIER FILING DATE: 1997-05-23
5	EARLIER APPLICATION NUMBER: 60/047,582
6	EARLIER FILING DATE: 1997-05-23
7	EARLIER APPLICATION NUMBER: 60/047,596
8	EARLIER FILING DATE: 1997-05-23
9	EARLIER APPLICATION NUMBER: 60/047,612
10	EARLIER FILING DATE: 1997-05-23
11	EARLIER APPLICATION NUMBER: 60/047,632
12	EARLIER FILING DATE: 1997-05-23
13	EARLIER APPLICATION NUMBER: 60/047,601
14	EARLIER FILING DATE: 1997-05-23
15	EARLIER APPLICATION NUMBER: 60/043,580
16	EARLIER FILING DATE: 1997-04-11
17	EARLIER APPLICATION NUMBER: 60/043,568
18	EARLIER FILING DATE: 1997-04-11
19	EARLIER APPLICATION NUMBER: 60/043,311
20	EARLIER FILING DATE: 1997-04-11
21	EARLIER APPLICATION NUMBER: 60/043,671
22	EARLIER FILING DATE: 1997-04-11
23	EARLIER APPLICATION NUMBER: 60/043,674
24	EARLIER FILING DATE: 1997-04-11
25	EARLIER APPLICATION NUMBER: 60/043,669
26	EARLIER FILING DATE: 1997-04-11
27	EARLIER APPLICATION NUMBER: 60/043,312
28	EARLIER FILING DATE: 1997-04-11
29	EARLIER APPLICATION NUMBER: 60/048,974
30	EARLIER FILING DATE: 1997-06-06
31	EARLIER APPLICATION NUMBER: 60/056,886
32	EARLIER FILING DATE: 1997-08-22
33	EARLIER APPLICATION NUMBER: 60/056,877
34	EARLIER FILING DATE: 1997-08-22
35	EARLIER APPLICATION NUMBER: 60/056,889
36	EARLIER FILING DATE: 1997-08-22
37	EARLIER APPLICATION NUMBER: 60/056,893
38	EARLIER FILING DATE: 1997-08-22
39	EARLIER APPLICATION NUMBER: 60/056,630
40	EARLIER FILING DATE: 1997-08-22
41	EARLIER APPLICATION NUMBER: 60/056,878
42	EARLIER FILING DATE: 1997-08-22
43	EARLIER APPLICATION NUMBER: 60/056,662
44	EARLIER FILING DATE: 1997-08-22
45	EARLIER APPLICATION NUMBER: 60/056,872
46	EARLIER FILING DATE: 1997-08-22
47	EARLIER APPLICATION NUMBER: 60/056,882
48	EARLIER FILING DATE: 1997-08-22
49	EARLIER APPLICATION NUMBER: 60/056,637
50	EARLIER FILING DATE: 1997-08-22
51	EARLIER APPLICATION NUMBER: 60/056,903
52	EARLIER FILING DATE: 1997-08-22
53	EARLIER APPLICATION NUMBER: 60/056,888
54	EARLIER FILING DATE: 1997-08-22
55	EARLIER APPLICATION NUMBER: 60/056,879
56	EARLIER FILING DATE: 1997-08-22
57	EARLIER APPLICATION NUMBER: 60/056,880
58	EARLIER FILING DATE: 1997-08-22
59	EARLIER APPLICATION NUMBER: 60/056,894
60	EARLIER FILING DATE: 1997-08-22
61	EARLIER APPLICATION NUMBER: 60/056,911
62	EARLIER FILING DATE: 1997-08-22
63	EARLIER APPLICATION NUMBER: 60/056,636

/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,874
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,910
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,864
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,631
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,845
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,892
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/057,761
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/047,595
/	EARLIER FILING DATE:	1997-05-23
/	EARLIER APPLICATION NUMBER:	60/047,599
/	EARLIER FILING DATE:	1997-05-23
/	EARLIER APPLICATION NUMBER:	60/047,588
/	EARLIER FILING DATE:	1997-05-23
/	EARLIER APPLICATION NUMBER:	60/047,585
/	EARLIER FILING DATE:	1997-05-23
/	EARLIER APPLICATION NUMBER:	60/047,594
/	EARLIER FILING DATE:	1997-05-23
/	EARLIER APPLICATION NUMBER:	60/047,589
/	EARLIER FILING DATE:	1997-05-23
/	EARLIER APPLICATION NUMBER:	60/047,590
/	EARLIER FILING DATE:	1997-05-23
/	EARLIER APPLICATION NUMBER:	60/047,614
/	EARLIER FILING DATE:	1997-05-23
/	EARLIER APPLICATION NUMBER:	60/043,578
/	EARLIER FILING DATE:	1997-04-11
/	EARLIER APPLICATION NUMBER:	60/043,576
/	EARLIER FILING DATE:	1997-04-11
/	EARLIER APPLICATION NUMBER:	60/047,501
/	EARLIER FILING DATE:	1997-05-23
/	EARLIER APPLICATION NUMBER:	60/043,670
/	EARLIER FILING DATE:	1997-04-11
/	EARLIER APPLICATION NUMBER:	60/056,632
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,664
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,876
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,881
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,909
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,875
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,862
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,887
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,908
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/048,964
/	EARLIER FILING DATE:	1997-06-06
/	EARLIER APPLICATION NUMBER:	60/057,650
/	EARLIER FILING DATE:	1997-09-05
/	EARLIER APPLICATION NUMBER:	60/056,884
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/057,669
/	EARLIER FILING DATE:	1997-09-05
/	EARLIER APPLICATION NUMBER:	60/049,610
/	EARLIER FILING DATE:	1997-06-13
/	EARLIER APPLICATION NUMBER:	60/061,060
/	EARLIER FILING DATE:	1997-10-02

US-08-943-731-200

Query Match 6.1%; Score 30.6; DB 4; Length 787;
Best Local Similarity 52.8%; Pred. No. 1.6;
Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 377 TCTGTGCTGTAGTGTGGCAGCTGCTGGTACCCGGTGGCTTGGAGAAAGTCAGCTCCCG 436
|||||
DB 340 TCTGGCTGAGGTGGGACCCAGCAGAGAGAGTGGCCGACGGTGCCTGGGACCC 281
|||||

QY 437 TCGTAGTGAGCACCTCTGGAACTCTCTCAGAGAGCACCCTTATTCGCCAAGTCCTTTT 496
|||||
DB 280 TGGGAGTGCCCACTCTGCGCAGTGGTGCTACCCACCCCTTACTTCCCGTGCTCT 221
|||||

QY 497 TGACA 501
|||

DB 220 GTGCA 216
||

RESULT 14

US-08-943-731-5/c

; Sequence 5, Application US/08943731

; Patent No. 6265157

; GENERAL INFORMATION:

; APPLICANT: PROCKOP, DARWIN J.

; APPLICANT: SPOTILA, LORETTA D.

; APPLICANT: DELTAS, CONSTANTINOS D.

; APPLICANT: SEREDA, LARISA

; APPLICANT: LARSON, ANDREA W.

; APPLICANT: PACK, MICHAEL

; APPLICANT: COLIGE, ALAIN

; APPLICANT: EARLY, JAMES

; APPLICANT: KORKKO, JARMO

; APPLICANT: ALA-KOKKO, LEENA, et al.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING

; TITLE OF INVENTION: ALTERED TYPE I OR TYPE IX COLLAGEN GENE SEQUENCES

; NUMBER OF SEQUENCES: 666

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.

; STREET: ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND

; CITY: PHILADELPHIA

; STATE: PA

; COUNTRY: USA

; ZIP: 19103-7086

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/943.731

; FILING DATE: 03-OCT-1997

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/212,322

; FILING DATE: 14-MAR-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/803,628

; FILING DATE: 03-DEC-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: DOYLE LEARY Ph.D., KATHRYN

; REGISTRATION NUMBER: 36,317

; REFERENCE/DOCKET NUMBER: 9598-27

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 215-965-1284

; TELEFAX: 215-567-2991

; TELEX: 831-494

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 20084 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

US-08-943-731-5

Query Match 6.1%; Score 30.6; DB 4; Length 20084;
Best Local Similarity 52.8%; Pred. No. 10;
Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 377 TCTGTGCTGTAGTGTGGCAGCTGCTGGTACCCGGTGGCTTGGAGAAAGTCAGCTCCCG 436
|||||
DB 14308 TCTGGCTGAGGTGGGACCCAGCAGAGAGAGTGGCCGACGGTGCCTGGGACCC 14249
|||||

QY 437 TCGTAGTGAGCACCTCTGGAACTCTCTCAGAGAGCACCCTTATTCGGCAAGTCCTTT 496
|||||
DB 14248 TGGAGTGCCCACTCTGCCAGTGGTGCTACCCACCCCTTACTTCCCGTGCTCT 14189
|||||

QY 497 TGACA 501
|||

DB 14188 GTGCA 14184
||

RESULT 15

US-09-734-675-3/c

; Sequence 3, Application US/09734675

; Patent No. 6365391

; GENERAL INFORMATION:

; APPLICANT: WESTER, MARION et al

; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND

; TITLE OF INVENTION: USES THEREOF

; FILE REFERENCE: CL000862

; CURRENT APPLICATION NUMBER: US/09/734,675

; CURRENT FILING DATE: 2000-12-13

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 38844

; TYPE: DNA

; ORGANISM: Human

US-09-734-675-3

Query Match 6.1%; Score 30.6; DB 4; Length 38844;
Best Local Similarity 46.5%; Pred. No. 15;
Matches 99; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 113 GCAGATCAACTAATTTTCATTTCAAGGCAGTTTTTAAAGAAAGTCTTGGAAAACAGACGCGG 172
|||||

DB 31579 GTATAAGATAAAACAAAACCAATGGGACATCTGGAATAAATCTAGAGACCTTAAACCG 31520
|||||

QY 173 CACCTTTCTCTAATCCAGCAAGTGATTCCTGCACACAGAGACAGAGAGTACAG 232
|||||

DB 31519 ATTCCTGCCCTATGTAGTGAAAGTTTATCTAGTTCTTAAATGGTAGACTGAGAATCAA 31460
|||||

QY 233 GATCAGTGGGTCTAAGTGTCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAAGAT 292
|||||

DB 31459 AAGCTTATCTTTATATTTCTTGATGTAATAAATGTTATAGAAATCTTCCAAATAGAT 31400
|||||

QY 293 TGAGTTTGCAATGTGAGTCTTTTGTCTTCTC 325
|||||

DB 31399 TTAGTCATCTCAGAAAAGCCCTTTTGCATTTTC 31367
|||||

Search completed: May 17, 2003, 16:47:42

Job time : 72.1908 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 15:56:54 ; Search time 1298.06 Seconds
(without alignments)
11232.524 Million cell updates/sec

Title: US-09-434-382-28_COPY_26164_26664

Perfect score: 501

Sequence: 1 99tatggagctgtgcgagg.....ttcgccaagtctttttgaca 501

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rtd.*
- 36: em.htg.man.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	501	100.0	118788	9	AC005277	AC005277 Homo sapi
2	303	60.5	2908	9	AF304370	AF304370 Homo sapi
3	301.4	60.2	2163	9	AK074244	AK074244 Homo sapi
4	301.4	60.2	2997	9	BC001939	BC001939 Homo sapi
5	301.4	60.2	3006	9	BC004158	BC004158 Homo sapi
6	299.4	59.8	2907	9	AF094687	AF094687 Homo sapi
7	296.6	59.2	2908	9	AF308698	AF308698 Pan trogl
8	296.6	59.2	2992	6	AX405825	AX405825 Sequence
9	294.6	58.8	2734	9	AK094333	AK094333 Homo sapi
10	294.6	58.8	2793	9	AK094012	AK094012 Homo sapi
11	294.6	58.8	2976	9	AK001392	AK001392 Homo sapi
C 12	286.4	57.2	429	11	G60268	G60268 SHGC-130964
13	254.6	50.8	2893	9	AF308694	AF308694 Gorilla g
14	158	31.5	386	6	AX069570	AX069570 Sequence
C 15	133.4	26.6	137	11	G42927	G42927 WIAF-176-ST
C 16	95.4	19.0	97	6	AX326639	AX326639 Sequence
C 17	95.4	19.0	97	6	AX326672	AX326672 Sequence
C 18	93.8	18.7	97	6	AX326638	AX326638 Sequence
19	93.8	18.7	97	6	AX326659	AX326659 Sequence
C 20	48.2	9.6	186836	2	AC097911	AC097911 Rattus no
21	42.2	8.4	34593	10	AF348157	AF348157 Mus muscu
22	42.2	8.4	195666	10	AL663045	AL663045 Mouse DNA
23	41.4	8.3	125020	9	AF429315	AF429315 Homo sapi
24	40.2	8.0	16402	4	MIBMCG	AF429315 Homo sapi
C 25	39.8	7.9	141115	2	AL450991	X72204 Balaenopter
26	39.8	7.9	154008	30	AC058799	AL450991 Homo sapi
27	39.8	7.9	187127	9	AC097359	AC058799 Homo sapi
C 28	37.8	7.5	125020	9	AF429315	AC097359 Homo sapi
29	37	7.4	338116	2	AC087159	AF429315 Homo sapi
C 30	36.8	7.3	159524	9	CNS01DUO	AC087159 Mus muscu
31	36.8	7.3	160276	2	AC013529	AL133305 Human chr
32	36.6	7.3	146638	9	AC010347	AC013529 Homo sapi
C 33	36.6	7.3	177902	9	AC026992	AC010347 Homo sapi
34	36.4	7.3	83675	9	AC004830	AC026992 Homo sapi
C 35	36.4	7.3	218542	2	AL731655	AC004830 Homo sapi
C 36	35.6	7.1	220989	2	AC105707	AL731655 Mus muscu
C 37	35.6	7.1	252576	2	AC111963	AC105707 Rattus no
C 38	35.4	7.1	105256	2	AC126180	AC111963 Rattus no
39	35.4	7.1	166386	2	AC126082	AC126180 Rattus no
C 40	35.4	7.1	212335	2	AC126941	AC126082 Rattus no
41	35.2	7.0	55965	9	AC104653	AC126941 Mus muscu
42	35.2	7.0	125685	2	AC069268	AC104653 Homo sapi
43	35.2	7.0	145465	2	AC069399	AC069268 Homo sapi
44	35.2	7.0	163496	9	AC092567	AC069399 Homo sapi
C 45	35	7.0	161687	2	AC117334	AC092567 Homo sapi
						AC117334 Rattus no

ALIGNMENTS

RESULT 1

AC005277 AC005277 118788 bp DNA linear PRI 23-JUL-1998

LOCUS Homo sapiens chromosome 17, clone hRPK.597_M_12, complete sequence.

DEFINITION AC005277

ACCESSION AC005277

VERSION AC005277.1 GI:3337311

KEYWORDS HTG.

SOURCE Homo sapiens.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 118788)

REFERENCE Birren,B., Fasman,K., Linton,L., Nusbaum,C. and Lander,E.

AUTHORS Homo sapiens chromosome 17, clone hRPK.597_M_12

TITLE Unpublished

JOURNAL

REFERENCE
AUTHORS

2 (bases 1 to 118788)
 Birren,B., Fasman,K., Linton,L., Nusbaum,C., Lander,E., Allen,N.,
 Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Boatn,C.,
 Boutwell,C., Brown,A., Castle,A., Cerny,J., Cooke,P., Depayre,E.,
 Devon,K., Dewar,K., Donelan,L., Etemadi,S., Ferreira,P.,
 FitzHugh,W., Forrest,C., Funke,R., Gage,D., Gardyna,S.,
 Gensheimer,S., Geraghty,K., Gilmartin,T., Grant,G., Hagos,B.,
 Harris,K., Horton,L., Howland,J.C., Hui,L., Jacotot,L., Kann,L.,
 Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
 Meidrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
 Nachman,A., Nahf,R., Naylor,J., Niloff,M., O'Connor,T., Pavlin,B.,
 Peterson,K., Riley,R., Roberts,D., Rossello,R., Roy,A., Shyam,R.,
 Stange-Thomann,N., Stiwell,J., Stojanovic,N., Stone,C.,
 Strickland,C., Subramanian,A., Torruella-Miller,I., Vassiliev,H.,
 Vo,A., Wagner,A., Wang,B., Wheeler,J., Wu,Y., Ye.W.J., Zhao,J. and
 Zody,M.

TITLE
JOURNAL

Direct Submission
 Submitted (10-JUL-1998) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
AUTHORS

3 (bases 1 to 118788)
 Birren,B., Fasman,K., Linton,L., Nusbaum,C., Lander,E., Allen,N.,
 Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Boatn,C.,
 Boutwell,C., Brown,A., Castle,A., Cerny,J., Cooke,P., Depayre,E.,
 Devon,K., Dewar,K., Donelan,L., Etemadi,S., Ferreira,P.,
 FitzHugh,W., Forrest,C., Funke,R., Gage,D., Gardyna,S.,
 Gensheimer,S., Geraghty,K., Gilmartin,T., Grant,G., Hagos,B.,
 Harris,K., Horton,L., Howland,J.C., Hui,L., Jacotot,L., Kann,L.,
 Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
 Meidrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
 Nachman,A., Nahf,R., Naylor,J., Niloff,M., O'Connor,T., Pavlin,B.,
 Peterson,K., Riley,R., Roberts,D., Rossello,R., Roy,A., Shyam,R.,
 Stange-Thomann,N., Stiwell,J., Stojanovic,N., Stone,C.,
 Strickland,C., Subramanian,A., Torruella-Miller,I., Vassiliev,H.,
 Vo,A., Wagner,A., Wang,B., Wheeler,J., Wu,Y., Ye.W.J., Zhao,J. and
 Zody,M.

TITLE
JOURNAL

Direct Submission
 Submitted (23-JUL-1998) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jul 23, 1998 this sequence version replaced gi:3335015.

All repeats were identified using RepeatMasker: Smit, A.F.A. &
 Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

FEATURES
source

Location/Qualifiers
 1..118788
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="17"
 /chromosome="17"
 /clone_lib="RPCI-II human BAC library"
 /clone="hRPK.597.M.12"
 795..884

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

/rpt_family="AluX"
 complement(4609..4895)
 /rpt_family="MER89"
 complement(4896..5186)
 /rpt_family="AluSp"
 5219..5479
 /rpt_family="L2"
 complement(5481..5740)
 /rpt_family="MER34"
 complement(5963..6275)
 /rpt_family="AluSp"
 complement(6411..6494)
 /rpt_family="MER5A"
 6830..7013
 /rpt_family="MLT1C"
 7014..7318
 /rpt_family="AluYb8"
 7319..7581
 /rpt_family="MLT1C"
 7721..8032
 /rpt_family="AluJb"
 8039..8333
 /rpt_family="AluSq"
 9431..9460
 /rpt_family="AT-rich"
 9662..9829
 /rpt_family="MIR"
 complement(11302..11728)
 /rpt_family="MER39"
 complement(11824..11884)
 /rpt_family="LTR16C"
 complement(11928..12108)
 /rpt_family="LTR16C"
 12453..12725
 /rpt_family="AluJb"
 13377..13785
 /rpt_family="MSTB"
 14118..14153
 /rpt_family="(CA)n"
 complement(15320..15620)
 /rpt_family="AluJo"
 15683..15759
 /rpt_family="(CA)n"
 complement(15838..15893)
 /rpt_family="MER91A"
 complement(15915..16080)
 /rpt_family="MER91A"
 complement(16474..16544)
 /rpt_family="L2"
 17907..17920
 /rpt_family="L1ME"
 18271..18311
 /rpt_family="L2"
 complement(18368..18529)
 /rpt_family="AluJb"
 complement(18530..18560)
 /rpt_family="(CA)n"
 complement(18561..18681)
 /rpt_family="AluJb"
 18760..19065
 /rpt_family="AluSg"
 complement(19638..19896)
 /rpt_family="MIR"
 complement(19971..20260)
 /rpt_family="L1ME3A"
 20263..20552
 /rpt_family="MER58B"
 complement(20553..20816)
 /rpt_family="L1ME3A"
 complement(20821..21069)
 /rpt_family="L1M4"
 complement(21071..21100)
 /rpt_family="(TAAA)n"

LIHETATLDGEEBAVEKTHSTTSOATISVGMNNAEFIMLNHFSORVAKVPLPSPNFS
EKVGVAEDHMKVCFGDEPTMPKLIPLKALFAGDIEEWEERREKRELQVRAALLSRE
LAGLEDEGPOKRAHTEEPQAKVRAQ"
BASE COUNT 702 a 793 c 826 g 587 t
ORIGIN

Query Match 60.5%; Score 303; DB 9; Length 2908;
Best Local Similarity 100.0%; Pred. No. 7.2e-93;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGTATGGAGCTGTGCGGAGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 60
Db 2606 GGTATGGAGCTGTGCGGAGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 2665
Qy 61 AGGACTGGTGTGGCAGACAGCCGGGCGGAGGCTGCCACACGGAAGCAAGCAGATGA 120
Db 2666 AGGACTGGTGTGGCAGACAGCCGGGCGGAGGCTGCCACACGGAAGCAAGCAGATGA 2725
Qy 121 ACTAATTTTCATTTCAAGGAGTTTTTAAAGAGTCTTGGAAACACAGCGGCGCACTTTC 180
Db 2726 ACTAATTTTCATTTCAAGGAGTTTTTAAAGAGTCTTGGAAACACAGCGGCGCACTTTC 2785
Qy 181 CTCTAATCCAGCAAAAGTGATTCCTCGCACACAGAGACAAAGCAGAGTAAACAGATCAGTG 240
Db 2786 CTCTAATCCAGCAAAAGTGATTCCTCGCACACAGAGACAAAGCAGAGTAAACAGATCAGTG 2845
Qy 241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAAGATTGAGTTTG 300
Db 2846 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAAGATTGAGTTTG 2905
Qy 301 CAA 303
Db 2906 CAA 2908

RESULT 3
AK074244
LOCUS AK074244 2163 bp mRNA linear PRI 15-FEB-2002
DEFINITION Homo sapiens cDNA FLJ23664 fis, clone HEP03495, highly similar to
Putative prostate cancer susceptibility protein.
ACCESSION AK074244
VERSION AK074244.1 GI:18676795
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens hepatoma cell_line:HePG2 cDNA to mRNA, clone_lib:HEP
clone:HEP03495.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2163)
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
Direct Submission
Submitted (14-FEB-2002) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure, Human Genome
Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology; cDNA library construction;
5'- & 3'-end and one pass sequencing; Department of Virology and Human
Genome Center, Institute of Medical Science, University of Tokyo
(partly supported by Science and Technology Agency).
LOCATION/Qualifiers
1. .2163
/organism="Homo sapiens"
/db_xref="taxon:9606"

/clone="HEP03495"
/cell_line="HePG2"
/cell_type="hepatoma"
/clone_lib="HEP"
/note="cloning vector: pME18SFL3"
BASE COUNT 530 a 582 c 613 g 438 t
ORIGIN

Query Match 60.2%; Score 301.4; DB 9; Length 2163;
Best Local Similarity 99.7%; Pred. No. 2.4e-92;
Matches 302; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GGTATGGAGCTGTGCGGAGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 60
Db 1845 GGTATGGAGCTGTGCGGAGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 1904
Qy 61 AGGACTGGTGTGGCAGACAGCCGGGCGGAGGCTGCCACACGGAAGCAAGCAGATGA 120
Db 1905 AGGACTGGTGTGGCAGACAGCCGGGCGGAGGCTGCCACACGGAAGCAAGCAGATGA 1964
Qy 121 ACTAATTTTCATTTCAAGGAGTTTTTAAAGAGTCTTGGAAACACAGCGGCGCACTTTC 180
Db 1965 ACTAATTTTCATTTCAAGGAGTTTTTAAAGAGTCTTGGAAACACAGCGGCGCACTTTC 2024
Qy 181 CTCTAATCCAGCAAAAGTGATTCCTCGCACACAGAGACAAAGCAGATCAGTG 240
Db 2025 CTCTAATCCAGCAAAAGTGATTCCTCGCACACAGAGACAAAGCAGATCAGTG 2084
Qy 241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAAGATTGAGTTTG 300
Db 2085 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAAGATTGAGTTTG 2144
Qy 301 CAA 303
Db 2145 CAA 2147

RESULT 4
BC001939
LOCUS BC001939 2997 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, putative prostate cancer susceptibility protein,
clone MGC:4102 IMAGE:2820640, mRNA, complete cds.
ACCESSION BC001939
VERSION BC001939.1 GI:12804972
KEYWORDS MGC.
SOURCE MGC.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2997)
Strausberg,R.
Direct Submission
Submitted (29-JAN-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Lilisa Prabhu, Parvaneh Saedi, Jacqueline
Schein, Duane Smalish, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.


```

Db 2735 AGGACTGGTGGCTGGCAGACCGCGGGCCAGAGGCTGCCACACCGAAGCAGAGATGA 2794
QY 121 ACTAATTTCAATTTCAAGGAGTTTTTAAAGAGTCTTTGGAACACAGACGGCGCACTTTTC 180
Db 2795 ACTAATTTCAATTTCAAGGAGTTTTTAAAGAGTCTTTGGAACACAGACGGCGCACTTTTC 2854
QY 181 CTCCTAATCCAGCAAAAGTATTCCTTCGCACACAGAGACAGAGATTAACAGGATCAGTG 240
Db 2855 CTCCTAATCCAGCAAAAGTATTCCTTCGCACACAGAGACAGAGATTAACAGGATCAGTG 2914
QY 241 GGTCTAAGTGTCCGAGACTTAAACGAAAATAGTATTTACGCTCAATAAAGATTGAGTTTG 300
Db 2915 GGTCTAAGTGTCCGAGACTTAAACGAAAATAGTATTTACGCTCAATAAAGATTGAGTTTG 2974
QY 301 CAA 303
Db 2975 CAA 2977

RESULT 6
AK094687
LOCUS AK094687.1 GI:21753794 mRNA linear PRI 15-JUL-2002
DEFINITION Homo sapiens cDNA FLJ37368 fis, clone BRAMY2024530, highly similar
to Homo sapiens prostate cancer susceptibility protein HPC2/ELAC2
mRNA.
ACCESSION AK094687
VERSION AK094687.1
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens amygdata cDNA to mRNA, clone lib:BRAMY2
clone:BRAMY2024530.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1
Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S.,
Fukuzumi,Y., Fujimori,Y., Komiyama,M., Sugiyama,T., Irie,R.,
Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,
Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,
Nagahara,K., Masuko,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2907)
Isogai,T. and Yamamoto,J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan (Fax:81-438-52-3986)
(E-mail:genomics@hri.co.jp, tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
Source
1..2907
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="BRAMY2024530"
/tissue_type="amygdala"
/clone_lib="BRAMY2"
/note="cloning vector: pME18SFL3"
BASE COUNT 696 a 788 c 839 g 584 t
ORIGIN
Query Match 59.8%; Score 299.4; DB 9; Length 2907;
Best Local Similarity 99.7%; Pred. No. 1.3e-91;
Matches 300; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 GGTATGAGCTGTGCCAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATCCCTCTT 60
Db 2607 GGTATGAGCTGTGCCAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATCCCTCTT 2666
QY 61 AGGACTGGTGGCTGGCAGACCGCGGGCCAGAGGCTGCCACACGGAAGCAAGCAGATGA 120
Db 2667 AGGACTGGTGGCTGGCAGACCGCGGGCCAGAGGCTGCCACACGGAAGCAAGCAGATGA 2726
QY 121 ACTAATTTCAATTTCAAGGAGTTTTTAAAGAGTCTTTGGAACACAGACGGCGCACTTTTC 180
Db 2727 ACTAATTTCAATTTCAAGGAGTTTTTAAAGAGTCTTTGGAACACAGACGGCGCACTTTTC 2786
QY 181 CTCTAATCCAGCAAAAGTATTCCTTCGCACACAGAGACAGAGATTAACAGGATCAGTG 240
Db 2787 CTCTAATCCAGCAAAAGTATTCCTTCGCACACAGAGACAGAGATTAACAGGATCAGTG 2846
QY 241 GGTCTAAGTGTCCGAGACTTAAACGAAAATAGTATTTACGCTGCAATAAAGATTGAGTTTG 300
Db 2847 GGTCTAAGTGTCCGAGACTTAAACGAAAATAGTATTTACGCTGCAATAAAGATTGAGTTTG 2906
QY 301 C 301
Db 2907 C 2907

RESULT 7
AF308698
LOCUS AF308698.1 GI:10946496 mRNA linear PRI 27-FEB-2001
DEFINITION Pan troglodytes ELAC2 mRNA, complete cds.
ACCESSION AF308698
VERSION AF308698.1
KEYWORDS Pan troglodytes.
SOURCE Pan troglodytes.
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Pan.
1 (bases 1 to 2908)
Tavtigian,S.V., Simard,J., Teng,D.H.F., Abtin,V., Baumgard,M.,
Beck,A., Camp,N.J., Carillo,A.R., Chen,Y., Dayananth,P.,
Desrochers,M., Dumont,M., Farnham,J.M., Frank,D., Frye,C.,
Ghaffari,S., Gupte,J.S., Hu,R., Iliev,D., Janecki,T., Kort,E.N.,
Laity,K.E., Leavitt,A., Leblanc,G., McArthur-Morrison,J.,
Pederson,A., Penn,B., Peterson,K.T., Reid,J.E., Richards,S.,
Schroeder,M., Smith,R., Snyder,S.C., Swedlund,B., Swensen,J.,
Thomas,A., Tranchant,M., Woodland,A.M., Labrie,F., Skolnick,M.H.,
Neuhausen,S., Rommens,J. and Cannon-Albright,L.A.
A candidate prostate cancer susceptibility gene at chromosome 17p
Nat. Genet. 27 (2), 172-180 (2001)
21096977
JOURNAL MEDLINE
11175785
2 (bases 1 to 2908)
Tavtigian,S.V., Simard,J., Teng,D.H.F., Baumgard,M., Dayananth,P.,
Desrochers,M., Dumont,M., Farnham,J.M., Frank,D., Frye,C.,
Ghaffari,S., Gupte,J.S., Hu,R., Iliev,D., Janecki,T., Kort,E.N.,
Laity,K., Leavitt,A., Leblanc,G., McArthur-Morrison,J.,
Pederson,A., Penn,B., Peterson,K.T., Reid,J.E., Richards,S.,
Schroeder,M., Smith,R., Snyder,S.C., Swedlund,B., Swensen,J.,
Thomas,A., Tranchant,M., Woodland,A.M., Labrie,F., Skolnick,M.H.,
Neuhausen,S., Rommens,J. and Cannon-Albright,L.A.
Pan troglodytes ortholog of human HPC2/ELAC2
Unpublished
3 (bases 1 to 2908)
Tavtigian,S.V., Simard,J., Teng,D.H.F., Baumgard,M., Dayananth,P.,
Desrochers,M., Dumont,M., Farnham,J.M., Frank,D., Frye,C.,
Ghaffari,S., Gupte,J.S., Hu,R., Iliev,D., Janecki,T., Kort,E.N.,
Laity,K., Leavitt,A., Leblanc,G., McArthur-Morrison,J.,
Pederson,A., Penn,B., Peterson,K.T., Reid,J.E., Richards,S.,
Schroeder,M., Smith,R., Snyder,S.C., Swedlund,B., Swensen,J.,
Thomas,A., Tranchant,M., Woodland,A.M., Labrie,F., Skolnick,M.H.,
Neuhausen,S., Rommens,J. and Cannon-Albright,L.A.
Direct Submission
Submitted (27-SEP-2000) Myriad Genetics, Inc., 320 Wakara Way, Salt

```

FEATURES	HYSEQ, INC. (US)	Location/Qualifiers	
source	1..2908	/organism="Homo sapiens"	
		/db_xref="taxon:9598"	
CDS	1..2481	/codon_start=1	
		/product="ELAC2"	
		/protein_id="GI:10946497"	
		/db_xref="GI:10946497"	
		/translation="MMALCSLLRSAAAGTMSQGRITISQAPARRERPRKDPRLHLRTRE KRGPSGSGGNTVYLOVVAAGSRDGAALVVFSEFNRYLFNCCEGVORLMOEHKLKV ARLDNIPLTRHMSNVGSLGSMILTKETGLPKVLSPGPPQLEKYLEAIKIFSPGLKG LELAVRPHSAFEYEDETMTVYQIPIHSEQRKGQWSPERPLSRSPERSSESSEN ENEPHPSQORRVDSSLVAFIKLHKGFLVLAKEMLGKLPVGTAAIAPIIA AVKDGKSTHEGRELAEELCTPPDGAAFVVECPDESFTQICENATFORYOQKAD APVALVHMAPSVLSDRYQOMERFGPDTQHLVLNENCASVHNLRSKHTQOTLNLI HPDIFPLLTSPCKKEGPTLSVPMVQCECLKYQLRPRERWQDAIITCNPEEFIEA LQLENFQOSVQERYRSADGPAPAEKRSQYPEIIFLTGSAIPMKIRNVSAITLVNISP DTSLLDCGEGTFCGLCRHYGDQVDRVLGTLAAVVFVSHLHADHTGUPSLIQRERL ASLGPFLHLLVAPNQLKAWLQYHNCQCEVLHHISMIIPAKCLOEGAEIISPAVERL ISLLRTCDLEEFOTCLVRHCKHAFGALVHTSGWKVYSGDTMPCEALVRMGKDATL LIHEATLEDGLEEBAVEKTHSTSOAISVGMNMAEFIMLNHFQRYAKVPLFSPNFS EKVGVAEDHMKVCFGRDTPMKLIPLKALFAGDIEEMERERREKRELQVRAALLSRE LAGGLEDEPQOKRAHTEEPQAKKVRQA"	
BASE COUNT	712 a	788 c	819 g 589 t
ORIGIN			
Query Match	59.2%;	Score 296.6;	DB 9; Length 2908;
Best Local Similarity	98.7%;	Pred. No. 1.2e-90;	
Matches	299;	Conservative	0; Mismatches 4; Indels 0; Gaps 0;
Qy	1	GGTATGGAGCTGTGCGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT	60
Db	2606	GGTATGGAGCTGTGCGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT	2665
Qy	61	AGGACTGTGCTGGCAGACGCGCGGCGAGAGGCTGCCACAGCAAGCAGAGATGA	120
Db	2666	AGGACTGTGCTGGCAGACGCGCGGCGAGAGGCTGCCACAGCAAGCAGAGATGA	2725
Qy	121	ACTAATTTTCATTTCAAGGAGCTTTTAAAGAAGCTTTTGAAGACAGACGCGCCTTTC	180
Db	2726	ACTAATTTTCATTTCAAGGAGCTTTTAAAGAAGCTTTTGAAGACAGACGCGCCTTTC	2785
Qy	181	CTCTAATCCAGCAAGTGTATCCCTGSCACACAGAGACAGCAGAGTAAACAGGATCAGTG	240
Db	2786	CTCTAATCCAGCAAGTGTATCCCTGSCACACAGAGACAGCAGAGTAAACAGGATCAGTG	2845
Qy	241	GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTCAGCTGCAATAAGATTGAGTTTG	300
Db	2846	GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTCAGCTGCAATAAGATTGAGTTTG	2905
Qy	301	CAA 303	
Db	2906	CAA 2908	
RESULT 8			
AX405825			
LOCUS	AX405825	2992 bp	DNA linear PAT 14-JUN-2002
DEFINITION	Sequence 240 from Patent WO0222660.		
ACCESSION	AX405825		
VERSION	AX405825.1	GI:21439089	
KEYWORDS	human.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE			
AUTHORS	1		
	Tang, Y. T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q. A., Ren, F.,		
	Xue, A. J., Yang, Y., Wehrman, T. and Drmanac, R. T.		
TITLE	Novel nucleic acids and polypeptides		
JOURNAL	Patent: WO 0222660-A 240 21-MAR-2002;		
FEATURES	HYSEQ, INC. (US)	Location/Qualifiers	
source	1..2992	/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
CDS	75..2555	/note="unnamed protein product"	
		/codon_start=1	
		/protein_id="CAD34932.1"	
		/db_xref="GI:21439090"	
		/translation="MMALCSLLRSAAAGTMSQGRITISQAPARRERPRKDPRLHLRTRE KRGPSGSGGNTVYLOVVAAGSRDGAALVVFSEFNRYLFNCCEGVORLMOEHKLKV ARLDNIPLTRHMSNVGSLGSMILTKETGLPKVLSPGPPQLEKYLEAIKIFSPGLKG LELAVRPHSAFEYEDETMTVYQIPIHSEQRKGQWSPERPLSRSPERSSESSEN ENEPHPSQORRVDSSLVAFIKLHKGFLVLAKEMLGKLPVGTAAIAPIIA AVKDGKSTHEGRELAEELCTPPDGAAFVVECPDESFTQICENATFORYOQKAD APVALVHMAPSVLSDRYQOMERFGPDTQHLVLNENCASVHNLRSKHTQOTLNLI HPDIFPLLTSPCKKEGPTLSVPMVQCECLKYQLRPRERWQDAIITCNPEEFIEA LQLENFQOSVQERYRSADGPAPAEKRSQYPEIIFLTGSAIPMKIRNVSAITLVNISP DTSLLDCGEGTFCGLCRHYGDQVDRVLGTLAAVVFVSHLHADHTGUPSLIQRERL ASLGPFLHLLVAPNQLKAWLQYHNCQCEVLHHISMIIPAKCLOEGAEIISPAVERL ISLLRTCDLEEFOTCLVRHCKHAFGALVHTSGWKVYSGDTMPCEALVRMGKDATL LIHEATLEDGLEEBAVEKTHSTSOAISVGMNMAEFIMLNHFQRYAKVPLFSPNFS EKVGVAEDHMKVCFGRDTPMKLIPLKALFAGDIEEMERERREKRELQVRAALLSRE LAGGLEDEPQOKRAHTEEPQAKKVRQA"	
BASE COUNT	725 a	807 c	859 g 601 t
ORIGIN			
Query Match	59.2%;	Score 296.6;	DB 6; Length 2992;
Best Local Similarity	98.7%;	Pred. No. 1.2e-90;	
Matches	299;	Conservative	0; Mismatches 4; Indels 0; Gaps 0;
Qy	1	GGTATGGAGCTGTGCGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT	60
Db	2680	GGTATGGAGCTGTGCGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT	2739
Qy	61	AGGACTGTGCTGGCAGACGCGCGGCGAGAGGCTGCCACAGCAAGCAGAGATGA	120
Db	2740	AGGACTGTGCTGGCAGACGCGCGGCGAGAGGCTGCCACAGCAAGCAGAGATGA	2799
Qy	121	ACTAATTTTCATTTCAAGGAGCTTTTAAAGAAGCTTTTGAAGACAGACGCGCCTTTC	180
Db	2800	ACTAATTTTCATTTCAAGGAGCTTTTAAAGAAGCTTTTGAAGACAGACGCGCCTTTC	2859
Qy	181	CTCTAATCCAGCAAGTGTATCCCTGSCACACAGAGACAGCAGAGTAAACAGGATCAGTG	240
Db	2860	CTCTAATCCAGCAAGTGTATCCCTGSCACACAGAGACAGCAGAGTAAACAGGATCAGTG	2919
Qy	241	GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTCAGCTGCAATAAGATTGAGTTTG	300
Db	2920	GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTCAGCTGCAATAAGATTGAGTTTG	2979
Qy	301	CAA 303	
Db	2980	CAA 2982	
RESULT 9			
AX094333			
LOCUS	AX094333	2734 bp	mrna linear PRI 15-JUL-2002
DEFINITION	Homo sapiens cDNA FLJ37014 fis, clone BRACE2010203, highly similar to Homo sapiens prostate cancer susceptibility protein HPC2/ELAC2 mrna.		
ACCESSION	AX094333		
VERSION	AX094333.1	GI:21753374	
KEYWORDS	oligo capping; fis (full insert sequence).		
SOURCE	Homo sapiens cerebellum cDNA to mRNA, clone lib:BRACE2 clone:BRACE2010203.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE			
AUTHORS	1		
	Nishi, T., Ota, T., Nakagawa, S., Senoh, A., Mizuguchi, H., Inagaki, H.,		

Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahara,K., Masuho,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2734)
Isogai,T. and Yamamoto,J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES

Location/Qualifiers
1..2734
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="BRACE2010203"
/tissue_type="cerebellum"
/clone_lib="BRACE2"
/note="cloning vector: pME18SFL3"

BASE COUNT 671 a 708 c 724 g 631 t

ORIGIN

Query Match 58.8%; Score 294.6; DB 9; Length 2734;
Best Local Similarity 98.7%; Pred. No. 5.7e-90;
Matches 297; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGTATGGAGCTGTGCGAGGCTTGGGCTCCACATAGCACTAGTCTATAGATGCTCTT 60
Db 2434 GGTATGGAGCTGTGCGAGGCTTGGGCTCCACATAGCACTAGTCTATAGATGCTCTT 2493
QY 61 AGGACTGTGCTGGCAGCGCGGCGGAGGCTGCCACACGGAAGCAAGCAGATGA 120
Db 2494 AGGACTGTGCTGGCAGCGCTGGGCGGAGGCTGCCACACGGAAGCAAGCAGATGA 2553
QY 121 ACTAATTTTCATTTCAAGCGAGTTTTTAAAGAGTCTTGAAACACAGCGCGGCACCTTTC 180
Db 2554 ACTAATTTTCATTTCAAGCGAGTTTTTAAAGAGTCTTGAAACACAGCGCGGCACCTTTC 2613
QY 181 CTCTAATCCAGCAAGTGTATCCCTGCGACACGAGAGCAAGCAGATGAACAGGATCAGTG 240
Db 2614 CTCTAATCCAGCAAGTGTATCCCTGCGACACGAGAGCAAGCAGATGAACAGGATCAGTG 2673
QY 241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAGATTGAGTTTG 300
Db 2674 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAGATTGAGTTTG 2733
QY 301 C 301
Db 2734 C 2734

RESULT 10

AK094012 AK094012 2793 bp mRNA linear PRI 15-JUL-2002
LOCUS Homo sapiens cDNA FLJ36693 fis, clone UTERU2008901, highly similar
DEFINITION to Homo sapiens putative prostate cancer susceptibility protein
HPC2/ELAC2 mRNA.

ACCESSION

AK094012
VERSION AK094012.1 GI:21752986
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens uterus cDNA to mRNA, clone_lib:UTERU2
clone:UTERU2008901.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

1 Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiya,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahara,K., Masuho,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2793)
Isogai,T. and Yamamoto,J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1..2793
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UTERU2008901"
/tissue_type="uterus"
/clone_lib="UTERU2"
/note="cloning vector: pME18SFL3"

BASE COUNT 610 a 754 c 806 g 623 t

ORIGIN

Query Match 58.8%; Score 294.6; DB 9; Length 2793;
Best Local Similarity 98.7%; Pred. No. 5.7e-90;
Matches 297; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGTATGGAGCTGTGCGAGGCTTGGGCTCCACATAGCACTAGTCTATAGATGCTCTT 60
Db 2493 GGTATGGAGCTGTGCGAGGCTTGGGCTCCACATAGCACTAGTCTATAGATGCTCTT 2552
QY 61 AGGACTGTGCTGGCAGCGCGGCGGAGGCTGCCACACGGAAGCAAGCAGATGA 120
Db 2553 AGGACTGTGCTGGCAGCGCTGGGCGGAGGCTGCCACACGGAAGCAAGCAGATGA 2612
QY 121 ACTAATTTTCATTTCAAGCGAGTTTTTAAAGAGTCTTGAAACACAGCGCGGCACCTTTC 180
Db 2613 ACTAATTTTCATTTCAAGCGAGTTTTTAAAGAGTCTTGAAACACAGCGCGGCACCTTTC 2672
QY 181 CTCTAATCCAGCAAGTGTATCCCTGCGACACGAGAGCAAGCAGATGAACAGGATCAGTG 240
Db 2673 CTCTAATCCAGCAAGTGTATCCCTGCGACACGAGAGCAAGCAGATGAACAGGATCAGTG 2732
QY 241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAGATTGAGTTTG 300
Db 2733 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAGATTGAGTTTG 2792

RESULT 11

AK001392

LOCUS

DEFINITION

ACCESSION

AK001392

AK001392

AK001392

AK001392

AK001392

AK001392

AK001392

AK001392

AK001392

AK001392

AK001392

AK001392.1 GI:7022621
oligo capping; fits (full insert sequence).
Homo sapiens teratocarcinoma cell line:NT2 cDNA to mRNA,
clone lib:NT2RP2 clone:NT2RP2000985.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Isogai,T., Ota,T., Hayaashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
Wagatsuma,M., Hosoi,T., Kaku,Y., Kodaira,H., Kondo,H.,
Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,
Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,
Nakamura,Y., Nagahashi,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2976)
Isogai,T. and Otsuki,T.
Direct Submission
Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing and clone selection;
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.
Location/Qualifiers
1. .2976
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2RP2000985"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP2"
/note="cloning vector: pME18SFL3-mRNA from NT2 neuronal
precursor cells after 2-weeks retinoic acid (RA)
induction."
75.2639
/note="unnamed protein product"
/codon_start=1
/protein_id="BAA91666.1"
/db_xref="GI:7022622"
/translation="MWALCSLLRSAGRTMSQRTISOAPARRRPRKDLRLHLRTRE
KRGSCSGGPNVTYLVVAGSRDGAALVFSEFNRYLNCGEVORLQEHKLK
ARLDNIFTRMWSNVGSLGMLTLKETGLPKCVLSPPQLEKYLEAIKIFSGPLK
IELAMRPHSAPEYEDETVYQIPHSEQRKQHPQSPERPLSRSPSSSEN
ENEPHGVQSRGVRDSSLVAFICKHLKRNFLVKAKEMGLPVGTAIPIIA
AVDKGSIHTEGREILABELCTPPDGAFFVVECPDESFTQICENATFORQCKAD
APVALVHMASVLVDSRYQOMERFGPDTHLVLENACSVHNLRSKHIOQLNLI
HPDIFPLTSRCKGPTLSVPMVQEGCLLKQLRPREWQDAILTCNPEFVEAL
QUNPQQSVETRRQAQGPAPAEKRSQPEIFLGTSAPIKRNVSATLVNISPD
TSLDDCEGTFGQCRHYGDQDVLGTLAAVFVSHLHADHTGLPSILIQERALA
SLGKFLPLLVAPNKLAWQYNNQVQLHSMIPAKCLQESISSPAVERLI
SSLRTCLDEEFQTCVLRCKHAFGCVLHVTSGWKVYSGDTPCEALVRMGKATLL
THEATLDGLEEEAVEKTHSTTSQAI SVGMNNAEFIMNHFSQRYAKVPLFSPNSE
KGVAFDHKVCGLGFPPTMPKLIPLKALFAGDIEEMERREKRELQVRAALLSREL
AGGLSDGSLRSRSGPTORPRRSPSPSDGLGDFELRLRCVCFPHAPVSALLAGRS"
BASE COUNT 712 a 807 c 856 g 601 t
ORIGIN
Query Match 58.8%; Score 294.6; DB 9; Length 2976;
Best Local Similarity 98.7%; Pred. No. 5.8e-90;
Matches 297; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GGTATGGAGCTGTGCCGAGCTTGGGCTCCACATAGCACTAGTCTATAGATGCCTCTT 60
|||||
Db 2676 GGTATGGAGCTGTGCCGAGCTTGGGCTCCACATAGCACTAGTCTATAGATGCCTCTT 2735
|||||
QY 61 AGGACTGGTCTGGCACACGCCCGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA 120

Db 2736 AGGACTGGTCTGGCACAGCTGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA 2795
|||||
QY 121 ACTAATTTTCATTTCAAGGCAGTCTTTTAAAGAAAGTCTTTGAAACAGACCGCGGCACCTTTTC 180
|||||
Db 2796 ACTAATTTTCATTTCAAGGCAGTCTTTTAAAGAAAGTCTTTGAAACAGACCGCGGCACCTTTTC 2855
|||||
QY 181 CTCTAATCAGCAAAAGTGTATTCCTTCGACACACAGAGACAAGCAGAGTAAACAGATCAGTG 240
|||||
Db 2856 CTCTAATCAGCAAAAGTGTATTCCTTCGACACACAGAGACAAGCAGAGTAAACAGATCAGTG 2915
|||||
QY 241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTCAGCTCAATAAAGATTGAGTTTG 300
|||||
Db 2916 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTCAGCTCAATAAAGATTGAGTTTG 2975
|||||
QY 301 C 301
Db 2976 C 2976
RESULT 12
G60268/c
LOCUS G60268 SHGC-130964 Human Homo sapiens STS genomic, linear STS 30-MAR-2000
DEFINITION G60268
ACCESSION G60268
VERSION G60268.1 GI:6125437
KEYWORDS STS.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 429)
AUTHORS Olivier,M. and Cox,D.R.
TITLE Unpublished, Olivier, M., Cox, D.R. (2000)
JOURNAL Unpublished (2000)
COMMENT Contact: Michael Olivier, David R. Cox
Stanford Human Genome Center
Stanford University School of Medicine
4005 Miranda Ave. 2nd Fl., Palo Alto, CA 94025, USA
Tel: (650) 320-5800
Fax: (650) 320-5801
Email: olivier@shgc.stanford.edu
Primer A: TTTGCTGGATTAGAGGAAAGGTG
Primer B: AGTGAAGATCTGGAGACCCCTGAA
STS size: 322
PCR Profile:
Initial incubation: 95 degrees C for 10 minutes.
Denaturation: 94 degrees C for 30 seconds
Annealing: 60 degrees C for 30 seconds
Polymerization: 72 degrees C for 23 seconds
PCR Cycles: 30
Thermal Cycler: Perkin Elmer 9700
Protocol:
Template: 25 ng
Primer: each 1 uM
dNTPs: each 200 uM
Ampliq Gold Polymerase: 0.07 units/ul
Total Vol: 5 ul
Buffer: MgCl2: 2.5 mM
KCl: 50 mM
Tris-HCl: 10 mM
pH: 8.3
Finished human sequence in NCBI. STSs designed and developed at the
Stanford Human Genome Center.
Location/Qualifiers
1. 429
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="17"
FEATURES
source

```
/clone.lib="Human"
STS
primer_bind 106..427
primer_bind 106..128
complement(405..427)
BASE COUNT 91 a 115 c 108 g 115 t
ORIGIN
Query Match 57.2%; Score 286.4; DB 11; Length 429;
Best Local Similarity 99.3%; Pred. No. 2.4e-87;
Matches 298; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1 GGTATGGAGCTGTGGCGAGGCTTGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 60
DB 299 GGTATGGAGCTGTGGCGAGGCTTGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 240
QY 61 AGGACTGTGCTGGCAGACCGCGCGGCGAGGCTGCCACACGAGCAAGCAAGCAGATGA 120
DB 239 AGGACTGTGCTGTGCACACCGCGCGGCGAGGCTGCCACACGAGCAAGCAGATGA 181
QY 121 ACTAATTTTCATTTCAAGGAGCTTTTAAAGAAGCTTGTGGAACACAGCGGCACCTTTC 180
DB 180 ACTAATTTTCATTTCAAGGAGCTTTTAAAGAAGCTTGTGGAACACAGCGGCACCTTTC 121
QY 181 CTCCTAATCAGCAAGTAGTATTCCTGCACACAGAGCAAGCAGAGTAGTACAGGATCAGTG 240
DB 120 CTCCTAATCAGCAAGTAGTATTCCTGCACACAGAGCAAGCAGAGTAGTACAGGATCAGTG 61
QY 241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTCAGTCTGCAATAAGATTGAGTTTG 300
DB 60 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTCAGTCTGCAATAAGATTGAGTTTG 1
RESULT 13
AF308694 2893 bp mRNA linear PRI 27-FEB-2001
LOCUS Gorilla gorilla ELAC2 (ELAC2) mRNA, complete cds.
DEFINITION Gorilla gorilla.
ACCESSION AF308694
VERSION AF308694.1 GI:10946488
KEYWORDS
SOURCE
ORGANISM
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Gorilla.
REFERENCE
1 (bases 1 to 2893)
Tavtigian,S.V., Simard,J., Teng,D.H.F., Abcin,V., Baumgard,M.,
Beck,A., Camp,N.J., Carillo,A.R., Chen,Y., Dayananth,P.,
Desrochers,M., Dumont,M., Farnham,J.M., Frank,D., Frye,C.,
Ghaffari,S., Gupte,J.S., Hu,R., Iliev,D., Janecki,T., Kort,E.N.,
Laity,K.E., Leavitt,A., Leblanc,G., McArthur-Morrison,J.,
Pederson,A., Penn,B., Peterson,K.T., Reid,J.E., Richards,S.,
Schroeder,M., Smith,R., Snyder,S.C., Swedlund,B., Swensen,J.,
Thomas,A., Tranchant,M., Woodland,A.M., Labrie,F., Skolnick,M.H.,
Neuhausen,S., Rommens,J. and Cannon-Albright,L.A.
A candidate prostate cancer susceptibility gene at chromosome 17p
Nat. Genet. 27 (2), 172-180 (2001)
21096977
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
2 (bases 1 to 2893)
Tavtigian,S.V., Simard,J., Teng,D.H.F., Baumgard,M., Beck,A.,
Camp,N.J., Carillo,A.R., Chen,Y., Dayananth,P., Desrochers,M.,
Dumont,M., Farnham,J.M., Frank,D., Frye,C., Ghaffari,S.,
Ghaffari,S., Hu,R., Iliev,D., Janecki,T., Kort,E.N., Laity,K.E.,
Leavitt,A., Leblanc,G., McArthur-Morrison,J., Pederson,A., Penn,B.,
Pederson,K.T., Reid,J.E., Richards,S., Schroeder,M., Smith,R.,
Snyder,S.C., Swedlund,B., Swensen,J., Thomas,A., Tranchant,M.,
Woodland,A.M., Labrie,F., Skolnick,M.H., Neuhausen,S., Rommens,J.
and Cannon-Albright,L.A.
Gorilla gorilla ortholog of human HPC2/ELAC2
Unpublished
3 (bases 1 to 2893)
Tavtigian,S.V., Simard,J., Teng,D.H.F., Baumgard,M., Beck,A.,
Camp,N.J., Carillo,A.R., Chen,Y., Dayananth,P., Desrochers,M.,
Dumont,M., Farnham,J.M., Frank,D., Frye,C., Ghaffari,S.,
Ghaffari,S., Hu,R., Iliev,D., Janecki,T., Kort,E.N., Laity,K.E.,
Leavitt,A., Leblanc,G., McArthur-Morrison,J., Pederson,A., Penn,B.,
Pederson,K.T., Reid,J.E., Richards,S., Schroeder,M., Smith,R.,
Snyder,S.C., Swedlund,B., Swensen,J., Thomas,A., Tranchant,M.,
Woodland,A.M., Labrie,F., Skolnick,M.H., Neuhausen,S., Rommens,J.
and Cannon-Albright,L.A.
Gorilla gorilla ortholog of human HPC2/ELAC2
Unpublished
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
Gupte,J.S., Hu,R., Iliev,D., Janecki,T., Kort,E.N., Laity,K.E.,
Leavitt,A., Leblanc,G., McArthur-Morrison,J., Pederson,A., Penn,B.,
Pederson,K.T., Reid,J.E., Richards,S., Schroeder,M., Smith,R.,
Snyder,S.C., Swedlund,B., Swensen,J., Thomas,A., Tranchant,M.,
Woodland,A.M., Labrie,F., Skolnick,M.H., Neuhausen,S., Rommens,J.
and Cannon-Albright,L.A.
Direct Submission
Submitted (27-SEP-2000) Myriad Genetics, Inc., 320 Wakara Way, Salt
Lake City, UT 84103, USA
Location/Qualifiers
source
1..2893
/organism="Gorilla gorilla"
/db_xref="taxon:9593"
gene
1..2893
/genes="ELAC2"
CDS
1..2481
/genes="ELAC2"
/notes="similar to Homo sapiens HPC2/ELAC2"
/codon_start=1
/product="ELAC2"
/protein_id="AAG24916.1"
/db_xref="GI:10946489"
/translat="MMALCSLLRSAAAGRTMSQGRITISQAPARRERPKDPLRLHRTRE
KRPGSCSGPNVTYLVQVVAAGSRDGAALYVFSEFNRYLFNCGEGVQRLMQEHLKV
VRJDNIFLTRMHSNVGLSGMILTILKETGLPKVLGSPPOLEKYLEAIKIFSPKLG
IELAVRPHSAPYEYEDTMYVQIPIHSEQRGRHQWQSPERPLSLSPSSDSSEN
ENEPHGHVSQRRGVRDSSLVAFIKLHKGKGNFLVLKAKEMGLPVGTAATAPIIA
AVDKGSIITHEGREILAEELCTPDFAAGVAVVECFDIFQICENATFORVGGKAD
APVALVHMVAPESVLDVSRVQMMERFGDTHLVNENCAVSHNRSHKIQOLNLI
HPDIFPLLTSPCKKEGPTLSVPMVOGECLLQYLRPRREQDAILITCNPEFIVEA
LQLPNFQSVQVRRSVQDVPAEKRQVPEIIFLGTSAI PMKIRNVATLVNISP
DTSLLDCCSEGTGQLCRHYGDQVDRVLGTAAVVFVSHLHADHTGLNLLQREOAL
ASUGKEHPLLVVAPSQLKAWLQQYHNQCEVLHHSMTIPAKLQGEAEISSAVERL
ISSLRITCDELEFQTCVLRCHAFQCALVHTSGWKVYSGDTPCEALVRMGKDATL
LIHEATLEDELEEEAEKHTSTTSQASVSGMRMAEFIMLNHFQRYKAYPLPSPNFN
EKGVAFDHMKVCFGDPPTMPKLIPLPKALFAGDIEEMERREKRELQVRAALLSGE
LAGGLEDEGEPEQKRAHTEBPQAKVRAQ"
BASE COUNT 705 a 787 c 815 g 586 t
ORIGIN
Query Match 50.8%; Score 254.6; DB 9; Length 2893;
Best Local Similarity 93.4%; Pred. No. 4.1e-76;
Matches 283; Conservative 0; Mismatches 4; Indels 16; Gaps 1;
QY 1 GGTATGGAGCTGTCCGAGGCTTGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 60
DB 2606 GGTATGGAGCTGTCCGAGGCTTGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 2655
QY 61 AGGACTGTGCTGGCAGACCGCGCGGCGAGGCTGCCACACGAGCAAGCAGATGA 120
DB 2656 -----GGTGTGCTGGCAGACCGCGCGGCGAGGCTGCCACACGAGCAAGCAGATGA 2709
QY 121 ACTAATTTTCATTTCAAGGAGCTTTTAAAGAAGCTTTTAAAGAAGCTTTTAAAGAAGCTTTT 180
DB 2710 ACTAATTTTCATTTCAAGGAGCTTTTAAAGAAGCTTTTAAAGAAGCTTTTAAAGAAGCTTTT 2769
QY 181 CTCCTAATCAGCAAGTAGTATTCCTGCACACAGAGCAAGCAGAGTAGTACAGGATCAGTG 240
DB 2770 CTCCTAATCAGCAAGTAGTATTCCTGCACACAGAGCAAGCAGAGTAGTACAGGATCAGTG 2829
QY 241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTCAGTCTGCAATAAGATTGAGTTTG 300
DB 2830 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTCAGTCTGCAATAAGATTGAGTTTG 2889
QY 301 CAA 303
DB 2890 CAA 2892
RESULT 14
AX069570 386 bp DNA linear PAT 25-JAN-2001
LOCUS AX069570
DEFINITION Sequence 42 from Patent WO0102568.
```


ACCESSION AX069570
VERSION AX069570.1 GI:12579355
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 386)
Williams, L.T., Escobedo, J., Innis, M.A., Garcia, P.D., Klinger, J.,
Kassam, A., Reinhard, C., Randazzo, F., Kennedy, G.C., Pot, D.,
Lamson, G., Drmanac, R., Crkjenjakov, F., Drmanac, S., Dickson, M.,
Labat, I., Leshkowitz, D., Kita, D., Garcia, V. and Strache-Crain, B.
Human genes and gene expression products
Patent: WO 0102568-A 42 11-JAN-2001;
JOURNAL CHIRON CORPORATION (US) ; HYSEQ. INC. (US)
FEATURES
Location/Qualifiers
1..386
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 82 a 90 c 111 g 103 t
ORIGIN
Query Match 31.5%; Score 158; DB 6; Length 386;
Best Local Similarity 100.0%; Pred. No. 6.3e-43;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 344 AGGGTCTGCTGTGCACACCTTGGAGAAGGCTCTCTGTGCTAGTGTGGCAGCTGCCTG 403
Db 7 AGGGTCTGCTGTGCACACCTTGGAGAAGGCTCTCTGTGCTAGTGTGGCAGCTGCCTG 66
QY 404 GTACCCGGTGGCTTGGAGAAGTACGTCCTCGTCTAGTGAGCACCTCTGGAACCTGTC 463
Db 67 GTACCCGGTGGCTTGGAGAAGTACGTCCTCGTCTAGTGAGCACCTCTGGAACCTGTC 126
QY 464 CTCAGAGAGCCACCTTATTCGCCAAGTCTTTTGACA 501
Db 127 CTCAGAGAGCCACCTTATTCGCCAAGTCTTTTGACA 164
RESULT 15
G42927/c
LOCUS G42927 137 bp mRNA linear STS 27-JAN-1999
DEFINITION WIAF-176-STS Human Thudson EST Homo sapiens STS cDNA, sequence
tagged site.
ACCESSION G42927
VERSION G42927.1 GI:4191844
KEYWORDS STS.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 137)
Wang, D.G., Fan, J.B., Siao, C.J., Berno, A., Young, P., Sapolsky, R.,
Chandour, G., Perkins, N., Winchester, E., Spencer, J., Kruglyak, L.,
Stein, L., Hsie, L., Topaloglou, T., Hubbell, E., Robinson, E.,
Mittmann, M., Morris, M.S., Shen, N., Kilburn, D., Rioux, J.,
Nusbaum, C., Rozen, S., Hudson, T.J., Lipshutz, R., Chee, M. and
Lander, E.S.
Large-scale identification, mapping, and genotyping of
single-nucleotide polymorphisms in the human genome
Science 280 (5366), 1077-1082 (1998)
98248615
PUBMED 9582121
COMMENT Synonyms: EST226740b, EST226740
Contact: Thomas Hudson
Whitehead Institute/MIT Center for Genome Research
Whitehead Institute for Biomedical Research
9 Cambridge Center, Cambridge MA 02142 USA
Tel: 617 252 1900
Fax: 617 252 1902
Email: thudson@genome.wi.mit.edu
Primer A: TATTGCAGCTGAATACTATTTCG
Primer B: AAGTCATGGAAACACGACGGC

STS size: 137
PCR Profile:
Presoak: 94 degrees C for 4.00 minutes
Denaturation: 94 degrees C for 50.0 seconds
Annealing: 58 degrees C for 1.50 minutes
Polymerization: 72 degrees C for 1.00 minutes
PCR Cycles: 30
Thermal Cycler: custom built by IAS, Costar, Cambridge MA
Protocol:
Template: 10 ng
Primer: each 5 pM
dNTPs: 4 nM
Taq Polymerase: 0.5 U
Total Vol: 20 uL
Buffer:
Mg2+: 1.5 mM
KCl: 50 mM
Tris-HCl: 10 mM
Gelatin: .001 %
Location/Qualifiers
1..137
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="36.40 cr from top of Chr17 linkage group"
/clone_lib="Human Thudson EST"
/note="STS derived from sequences in dbEST and the
Unigene collection."
STS
primer_bind 1..137
primer_bind complement(118..137)
BASE COUNT 28 a 30 c 31 g 46 t 2 others
ORIGIN
Query Match 26.6%; Score 133.4; DB 11; Length 137;
Best Local Similarity 97.8%; Pred. No. 1.5e-34;
Matches 134; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 151 AAGTCTTGGAAACACAGCGCGGCACCTTTCTCTCTAATCCAGCAAAGTGATTCCTGCACA 210
Db 137 AAGTCATGGAAACACAGCGCGGCACCTTTCTCTCTAATCCAGCAAANTGATTCCTGCANA 78
QY 211 CCAGAGACAAGCAGAGTAAACAGGATCAGTGGGTCTAAGTGTCCGAGACTTAACGAAAATA 270
Db 77 CCAGAGACAAGCAGAGTAAACAGGATCAGTGGGTCTAAGTGTCCGAGACTTAACGAAAATA 18
QY 271 GTATTTTCAGCTGCAATA 287
Db 17 GTATTTTCAGCTGCAATA 1
Search completed: May 17, 2003, 21:41:54
Job time : 1383.06 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 15:56:54 ; Search time 680.481 Seconds
(without alignments)

11900.029 Million cell updates/sec

Title: US-09-434-382-28_COPY_1_500

Perfect score: 500

Sequence: 1 taccagtgactgaattcta.....aagctctgagggaactgacgt 500

Scoring table:

IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_mam:*

24: em_gss_mus:*

25: em_gss_Other:*

26: em_gss_pro:*

27: em_gss_rod:*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	122.4	24.5	992	12	BF244530 601862835
2	119.6	23.9	546	13	BG943515 ax38g05.x
3	118.2	23.6	669	17	AG177687 Pan trogl
4	118.2	23.6	811	17	AQ742365 HS_5382B
C 5	117.8	23.6	399	17	AQ587593 CITBI-E1
C 6	117.8	23.6	450	17	AQ587553 CITBI-E1

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

7	117.8	23.6	727	9	AU121896
C 8	117.8	23.6	770	17	B02308
C 9	117.6	23.5	415	9	AI609972
C 10	117.6	23.5	669	17	AQ196057
C 11	117.4	23.5	389	17	AQ071874 HS_3018A
C 12	116.8	23.4	654	17	AG084157 Pan trogl
C 13	116.6	23.3	312	17	AQ230437 HS_2034A
C 14	116.4	23.3	389	17	B95260
C 15	116.4	23.3	415	17	AQ0628850
C 16	116.4	23.3	450	17	AQ086239 HS_2171B
C 17	116	23.2	527	17	AQ183714 HS_3205B
C 18	116	23.2	664	17	AQ343449
C 19	116	23.2	801	17	AQ740452 HS_5502A
C 20	115.8	23.1	523	17	AQ677233 HS_5522A
C 21	115.6	23.1	398	9	AL696113
C 22	115.6	23.1	500	17	B44892
C 23	115.4	23.1	324	9	AL707132
C 24	115.4	23.1	370	17	AZ756633
C 25	115.4	23.1	439	17	AQ378798
C 26	115.2	23.0	383	9	AA584125
C 27	115.2	23.0	426	9	AL043719
C 28	115.2	23.0	1080	13	BM558580
C 29	115	23.0	421	14	BQ614109
C 30	115	23.0	458	12	BF913236
C 31	114.8	23.0	390	17	AQ096225
C 32	114.6	22.9	670	17	AG013922
C 33	114.4	22.9	339	9	AI820920
C 34	114.4	22.9	426	12	BF589079
C 35	114.4	22.9	443	9	AI821670
C 36	114.4	22.9	450	17	AQ144748
C 37	114.4	22.9	459	9	AA478209
C 38	114.4	22.9	515	17	AQ019249
C 39	114.4	22.9	586	17	AQ583832
C 40	114.4	22.9	667	17	AQ059714
C 41	114.4	22.9	896	12	BF342636
C 42	114.2	22.8	516	10	BE148994
C 43	114.2	22.8	579	14	BM990962
C 44	114.2	22.8	600	10	AV759518
C 45	114.2	22.8	691	17	AG047439

ALIGNMENTS

RESULT 1
BF244530/c
LOCUS BF244530 992 bp mRNA linear EST 14-NOV-2000
DEFINITION 601862835f1 NIH_MGC_57 Homo sapiens CDNA clone IMAGE:4080393 5', mRNA sequence.
ACCESSION BF244530
VERSION BF244530.1 GI:11158461
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 992)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONETECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LiCM940 row: f column: 10
High quality sequence stop: 328.
Location/Qualifiers 1..992

FEATURES
source


```
source
1. .399
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2643H8"
/sex="male"
/cell_type="sperm"
/notes="Vector: pBelOBAC11; Site_1: EcoRI; Site_2: EcoRI;
CalTech Human BAC Library D"
BASE COUNT 119 a 83 c 108 g 89 t
ORIGIN

Query Match 23.6%; Score 117.8; DB 17; Length 399;
Best Local Similarity 73.8%; Pred. No. 1e-12;
Matches 163; Conservative 0; Mismatches 57; Indels 1; Gaps 1;

Qy 125 CTCCTGCTTTAGAGCTTGCTCTCTATTCTTCTTTTCCAAAACACCTACAAATTTT 184
Db 399 CTCCTCCCTTCAAATATCAACATTTCTACAATAGAGGCATACAGAGATGACAACCTTTT 340

Qy 185 TGTCTTTGTTTGTGTTTGTGAGACAGGGTCTCGAGGTGTCACCCAGGCTGGAGTGCAG 244
Db 339 TTTCTCTCCCTTCAAATATCAACATTTCTACAATAGAGGCATACAGAGATGACAACCTTTT 280

Qy 245 TGGCGCGATTTCGACTACCGCAACCTCGGCTCCGCGCTT-AAGCGATTCTCTGCGCTC 303
Db 279 TGGCGCGATCTCGGCTCACTGCAAGTCCGCTCCAGGTTACGCCATTCTCTGCGCTC 220

Qy 304 AGCTCTCCCAAGTAGCTGGGACTACAAAGTTCGGGACACCAACG 344
Db 219 AGCTCTCCCAAGTAGCTGGGACTACAGGCACCGCCGCCCATG 179

RESULT 6
AQ587553/c
LOCUS
DEFINITION AQ587553 450 bp DNA linear GSS 07-JUN-1999
CITBI-EI-2643H14. TR CITBI-EI Homo sapiens genomic clone 2643H14,
DNA sequence.
ACCESSION AQ587553
VERSION AQ587553.1 GI:5014233
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (Bases 1 to 450)
AUTHORS Zhao, S., Adams, M.D., Nierman, W., Malek, J., Shizuoya, H., Simon, M. and
Venter, J.C.
TITLE Use of BAC End Sequences from CalTech Libraries for Sequence-Ready
Map Building
JOURNAL Unpublished (1997)
COMMENT Other GSSs: CITBI-EI-2643H14.TF
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES
source
1. .450
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2643H14"
/sex="male"
/cell_type="sperm"
/notes="Vector: pBelOBAC11; Site_1: EcoRI; Site_2: EcoRI;
CalTech Human BAC Library D"
BASE COUNT 119 a 83 c 108 g 89 t
ORIGIN

Query Match 23.6%; Score 117.8; DB 9; Length 727;
Best Local Similarity 73.8%; Pred. No. 8.6e-13;
Matches 163; Conservative 0; Mismatches 57; Indels 1; Gaps 1;

Qy 125 CTCCTGCTTTAGAGCTTGCTCTCTATTCTTCTTTTCCAAAACACCTACAAATTTT 184
Db 312 CTTTCTGTTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 371
```

```
source
1. .727
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MAMMA1001220"
/clone_lib="MAMMA1"
/tissue_type="mammary gland"
/notes="Vector: pWE18SFJ3"
BASE COUNT 170 a 159 c 146 g 248 t
ORIGIN

Query Match 23.6%; Score 117.8; DB 9; Length 727;
Best Local Similarity 73.8%; Pred. No. 8.6e-13;
Matches 163; Conservative 0; Mismatches 57; Indels 1; Gaps 1;

Qy 125 CTCCTGCTTTAGAGCTTGCTCTCTATTCTTCTTTTCCAAAACACCTACAAATTTT 184
Db 312 CTTTCTGTTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 371
```

```

QY 185 TGTTTTTTTGTGTTTGTGAGACAGGGTCTCGAGGTGTACCCAGGCTGGAGTGCAG 244
Db 372 TGTGTGTTTGTGTTTGTGTTGAGACAGAGTCTCTCTGTCTACCCAGGCTGGAGTGCAG 431

QY 245 TGGCGGATTTCCAGTACCGCAACTCCGCTCCCGCTT-AAAGGATTCCTCGCTC 303
Db 432 TGACGGATCTCGGCTCACTGCAAGTCTCGCTCCAGGTTACGCGGTTTCTTGCCTC 491

QY 304 AGCTCCCAAGTAGCTGGGACTACAACTCGGACACACCAG 344
Db 492 AGCTCCCAAGTAGCTGGGACTACAGGCAACCGCACACCAG 532

RESULT 8
LOCUS B02308/c
DEFINITION CSRL-151D1-u CSRL flow sorted Chromosome 11 specific cosmid Homo sapiens genomic clone cSRL-151D1, DNA sequence.
ACCESSION B02308
VERSION B02308.1 GI:1411586
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 770)
AUTHORS Jones, G.A., Burbee, D., Davies, C., Hahner, L., Oliver, T., Gilbert, M., Ward, T., Gillilan, E., Schagemann, J., Probst, S., Harris, J., DeFord, J., McFarland, J., Bursinski, K., Khan, M., Kupfer, K. and Garner, H.R.
TITLE Genomic Sequence Sampled Map of Chromosome 11
JOURNAL Unpublished (1996)
COMMENT Contact: Evans GA, Shane Probst
McDermott Center for Human Growth and Development
University of Texas Southwestern Medical Center At Dallas
5323 Harry Hines Blvd, Dallas TX 75235-8591
Tel: 214-648-1600
Fax: 214-648-1666
Email: gevang@utsw.swmed.edu, shane@mcdermott.swmed.edu
PCR Primers
FORWARD: ATAAAGGAGGTTCTCTAAAG
BACKWARD: CTTTCTCTACGAGGAGGAC
Seq primer: 77
Class: cosmid ends
High quality sequence stop: 770.
Location/Qualifiers
1..770
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="cSRL-151D1"
/cosmid="cSRL flow sorted Chromosome 11 specific cosmid"
/sex="female"
/cell_type="chimeric hamster somatic cell hybrid"
/notes="Vector: scos-1; Human Chromosome 11 specific cosmid library prepared from flow sorted human Chromosome 11 derived from Chinese Hamster Ovary (CHO) monochromosomal somatic cell hybrid, J1"
BASE COUNT 219 a 169 c 166 g 185 t 31 others
ORIGIN
Query Match 23.6%; Score 117.8; DB 17; Length 770;
Best Local Similarity 83.2%; Pred. No. 8.4e-13;
Matches 134; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 184 TTGTTTGTGTTTGTGTTTGTGAGACAGGTTCTCGAGTGTACCCAGGCTGGAGTGCAG 243
Db 270 TTTCTTTTGTGTTTGTGTTTGTGAGACAGTCTCGCTCTCGCGCAGGCTGGAGTGCAG 211

QY 244 GTGGCGGATTTCCAGTACCGCAACTCCGCTCCGCTTAAAGCGATTCCTCGCTC 303
Db 210 GTGGCGGATTCGGGTCACTCGAACCTCTGCTCGAGTTCAAGCGATTCCTCGCTC 151

```

```

QY 304 AGCTCCCAAGTAGCTGGGACTACAACTCGGACACACCAG 344
Db 150 AGCTCTCTAGTAGTGGATTACAGGCGCGCCACACCAG 110

RESULT 9
LOCUS A1609972/c
DEFINITION A1609972 Homo sapiens cdna clone IMAGE:224698 3, similar to contains Alu repetitive element; contains element MER9 repetitive element ;, mRNA sequence.
ACCESSION A1609972
VERSION A1609972.1 GI:4619139
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 415)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Herbert Morse, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco.
FEATURES
Source
1..415
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:224698"
/clone_lib="NCI CGAP HSC3"
/tissue_type="CD34+ T negative, patient with chronic myelogenous leukemia"
/lab_host="DH10B"
/notes="Organ: bone marrow; Vector: pAMP1; mRNA made from lymphoid tissue, cdna made by oligo-dt priming. Directionally cloned. Size-selected on agarose gel, average insert size 500 bp. Primary library, non-amplified. cDNA Library Preparation: David B. Krizman, Ph.D. Reference: Krizman et al. (1996) Cancer Research 56:5380-5383."
BASE COUNT 132 a 81 c 125 g 77 t
ORIGIN
Query Match 23.5%; Score 117.6; DB 9; Length 415;
Best Local Similarity 80.2; Pred. No. 1.1e-12;
Matches 138; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 172 ACACATACAAATTTGTTTGTGTTTGTGTTTGTGAGACAGGTTCTCGAGTGTACACCA 231
Db 412 ATACGACTTTTTTTTTTTTTTTTTTTTTTTTTTTTGTGAGACAGAGTCTCGCTGTTTCATCCA 353

QY 232 GGCTGGAGTGCAGTGGCGGATTTTCGACTCACCGCAACTCCGCTCCGCGTTAAGCGA 291
Db 352 GGCTGGAGTGCAGTGGTGTGATCTCGGCTCACTGCAACTCCGCTCCAGGTTCAAGCA 293

QY 292 TTCTCTCGCTCAGCTCCCAAGTAGCTGGGACTACAAAGCTGGGACACCAC 343
Db 292 TTCTCTCGCTCAGCTTCCCAAGTAGCTGGGATTACAGGCGCATGCCACCAC 241

RESULT 10
LOCUS AQ196057/c
SOURCE AQ196057
669 bp DNA linear GSS 20-APR-1999

```


GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 17:27:39 ; Search time 63.3902 Seconds
(without alignments)

10174.595 Million cell updates/sec

Title: US-09-434-382-28_COPY_1_500

Perfect score: 500

Sequence: 1 tatcagtgactgaattctta.....aagctctgagggaactgacgt 500

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 810007 seqs, 644960991 residues

Total number of hits satisfying chosen parameters: 1620014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	500	100.0	26664	9	US-09-988-626-28 Sequence 28, Appl
2	500	100.0	26664	9	US-09-988-687-28 Sequence 28, Appl
3	120.4	24.1	7680	10	US-09-939-581A-3 Sequence 3, Appl
4	120	24.0	1503841	9	US-09-946-807-1 Sequence 1, Appl
5	120	24.0	1503841	10	US-09-795-668-1 Sequence 1, Appl
6	120	24.0	1503841	10	US-09-795-686-1 Sequence 1, Appl
7	119.8	24.0	4045	9	US-09-764-891-8718 Sequence 8718, Ap
8	119.8	24.0	4045	9	US-09-764-891-8719 Sequence 8719, Ap
C 9	119.6	23.9	1160	8	US-09-927-939-78 Sequence 78, Appl
C 10	119.6	23.9	1160	9	US-09-954-531-146 Sequence 146, App
C 11	118.8	23.8	73467	9	US-10-237-859-3 Sequence 3, Appl
C 12	118.4	23.7	17752	10	US-09-748-127-3 Sequence 3, Appl
C 13	118.4	23.7	465237	10	US-09-933-267A-1 Sequence 1, Appl
C 14	118.2	23.6	5159	10	US-09-764-877-3707 Sequence 3707, Ap
C 15	118.2	23.6	6834	9	US-09-764-891-8002 Sequence 8002, Ap
C 16	118.2	23.6	6834	9	US-10-091-438-263 Sequence 263, App
C 17	118	23.6	1115	9	US-09-764-872-943 Sequence 943, App
C 18	118	23.6	1115	9	US-09-764-872-944 Sequence 944, App
C 19	117.6	23.5	6186	9	US-10-074-095-1075 Sequence 1075, Ap

20	117.6	23.5	6186	10	US-09-764-860-1075	Sequence 1075, Ap
C 21	117.2	23.4	433	9	US-09-946-807-1404	Sequence 1404, Ap
C 22	117.2	23.4	433	10	US-09-795-668-1404	Sequence 1404, Ap
C 23	117.2	23.4	433	10	US-09-795-668-1404	Sequence 1404, Ap
24	116.8	23.4	2295	9	US-09-764-891-7825	Sequence 7825, Ap
25	116.4	23.3	5426	10	US-09-798-029-7	Sequence 7, Appli
26	116	23.2	6191	9	US-10-074-095-1076	Sequence 1076, Ap
27	116	23.2	6191	9	US-10-074-095-1077	Sequence 1077, Ap
28	116	23.2	6191	10	US-09-764-860-1076	Sequence 1076, Ap
29	116	23.2	6191	10	US-09-764-860-1077	Sequence 1077, Ap
30	116	23.2	20444	9	US-09-764-891-9422	Sequence 9422, Ap
31	115.8	23.2	8220	10	US-09-797-908-3	Sequence 3, Appli
C 32	115.6	23.1	312	9	US-10-102-627-86	Sequence 86, Appl
C 33	115.6	23.1	312	9	US-10-102-627-90	Sequence 90, Appl
C 34	115.6	23.1	312	10	US-09-764-877-3132	Sequence 3132, Ap
C 35	115.6	23.1	6145	9	US-10-102-627-87	Sequence 87, Appl
C 36	115.6	23.1	6145	9	US-10-102-627-91	Sequence 91, Appl
37	115.6	23.1	6145	10	US-09-764-877-3133	Sequence 3133, Ap
C 38	115	23.0	401	9	US-09-946-807-326	Sequence 326, App
C 39	115	23.0	401	9	US-09-946-807-327	Sequence 327, App
C 40	115	23.0	401	10	US-09-795-668-326	Sequence 326, App
C 41	115	23.0	401	10	US-09-795-668-327	Sequence 327, App
C 42	115	23.0	401	10	US-09-795-686-326	Sequence 326, App
C 43	115	23.0	401	10	US-09-795-686-327	Sequence 327, App
44	115	23.0	1954	9	US-10-091-572-715	Sequence 715, App
45	115	23.0	1954	9	US-09-764-891-9139	Sequence 9139, Ap

ALIGNMENTS

RESULT 1

US-09-988-626-28
Sequence 28, Application US/09988626
Publication No. US20030044959A1
GENERAL INFORMATION:
APPLICANT: Tavtigian, Sean V.
APPLICANT: Teng, David H.F.
APPLICANT: Simard, Jacques
APPLICANT: Rommens, Johanna M.
APPLICANT: Myriad Genetics, Inc.
TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
FILE REFERENCE: 2318-258
CURRENT APPLICATION NUMBER: US/09/988,626
CURRENT FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: 09/564,805
PRIOR FILING DATE: 2000-05-05
PRIOR APPLICATION NUMBER: US 60/107,468
PRIOR FILING DATE: 1998-11-06
PRIOR APPLICATION NUMBER: 09/434,382
PRIOR FILING DATE: 1999-11-05
NUMBER OF SEQ ID NOS: 240
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 28
LENGTH: 26664
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (910)..(13104)
OTHER INFORMATION: exon 1: 910-1154; exon 2: 1736-1786; exon 3:
OTHER INFORMATION: 1925-1995; exon 4: 3025-3089; exon 5: 4361-4418;
OTHER INFORMATION: exon 6: 5582-5650; exon 7: 7075-7194; exon 8:
OTHER INFORMATION: 8186-8244; exon 9: 12878-12936; exon 10:
OTHER INFORMATION: 13032-13104;
NAME/KEY: misc feature
LOCATION: (13756)..(22917)
OTHER INFORMATION: exon 11: 13756-13868; exon 12: 15283-15378; exon
OTHER INFORMATION: 13: 16278-16416; exon 14: 16498-16583; exon 15:
OTHER INFORMATION: 19583-18701; exon 16: 20349-20445; exon 17:
OTHER INFORMATION: 22172-22310; exon 18: 22879-22917
NAME/KEY: misc feature


```
; LENGTH: 4045
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-8718

Query Match          24.0%; Score 119.8; DB 9; Length 4045;
Best Local Similarity 73.0%; Pred. No. 9.9e-24;
Matches 154; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

Qy 133 TTAGAGCTTGCTCTATTCTTCTTTCTTCCAAAAAACAACACTACAATTTTGTGTTGT 192
Db 1603 TTCTACATTGGCTTCTTTCTCGTTCCTTCTCTTCCCTTCTTCTCTCTCTC 1662

Qy 193 TTTGTTTTTTTGTAGACAGAGGCTCGAGGTGTACCCAGGCTGGAGTGCAGTGGCGCA 252
Db 1663 TCTCTTTCTTTTGTAGACAGAGTCTCACTCTGTGCGCCAGGCTGGAGTGCAGTGGCACGA 1722

Qy 253 TTTTCGACTCACCGCAACCTCCGCTCCGGCTTAAGCGATTCTCCNGCTCAGCCTCCCA 312
Db 1723 TCTCAGCTACTGCAACCTCTGCTCTGAGTTCAGCGATTCTCTGCTCAGCCTCCCA 1782

Qy 313 AGTAGCTGGGACTACAAGCTCGGGACACCAC 343
Db 1783 AGTAGCTGGGACTACAAGCATGGCCACCAC 1813

RESULT 8
US-09-764-891-8719
; Sequence 8719, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8719
; LENGTH: 4045
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-8719

Query Match          24.0%; Score 119.8; DB 9; Length 4045;
Best Local Similarity 73.0%; Pred. No. 9.9e-24;
Matches 154; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

Qy 133 TTAGAGCTTGCTCTATTCTTCTTCTTCTTCCAAAAAACAACACTACAATTTTGTGTTGT 192
Db 1603 TTCTACATTGGCTTCTTTCTCGTTCCTTCTCTTCCCTTCTTCTCTCTC 1662

Qy 193 TTTGTTTTTTTGTAGACAGAGGCTCGAGGTGTACCCAGGCTGGAGTGCAGTGGCGCA 252
Db 1663 TCTCTTTCTTTTGTAGACAGAGTCTCACTCTGTGCGCCAGGCTGGAGTGCAGTGGCACGA 1722

Qy 253 TTTTCGACTCACCGCAACCTCCGCTCCGGCTTAAGCGATTCTCCNGCTCAGCCTCCCA 312
Db 1723 TCTCAGCTACTGCAACCTCTGCTCTGAGTTCAGCGATTCTCTGCTCAGCCTCCCA 1782

Qy 313 AGTAGCTGGGACTACAAGCTCGGGACACCAC 343
Db 1783 AGTAGCTGGGACTACAAGCATGGCCACCAC 1813
```

```
RESULT 9
US-08-927-939-78/c
; Sequence 78, Application US/08927939
; Patent No. US20010006640A1
; GENERAL INFORMATION:
; APPLICANT: Grainger, David J.
; APPLICANT: Tatalick, Lauen Marie
```

```
; TITLE OF INVENTION: Compounds and methods to inhibit or
; TITLE OF INVENTION: augment an inflammatory response.
; FILE REFERENCE: 295.022US1
; CURRENT APPLICATION NUMBER: US/08/927,939
; CURRENT FILING DATE: 1997-09-11
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: PastSeq for Windows Version 3.0
; SEQ ID NO 78
; LENGTH: 1160
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (27)...(299)
US-08-927-939-78

Query Match          23.9%; Score 119.6; DB 8; Length 1160;
Best Local Similarity 82.5%; Pred. No. 5.6e-24;
Matches 137; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 179 AATTTTGTGTTTGTGTTTGTGTTTGTGAGACAGGTTCTCGAGGTGTCAACCCAGGCTGGA 238
Db 767 ATTATTTTGTGTTTGTGTTTGTGTTTGTGAGACGAGTCTCGCTCTGTGCGCCAGGCTGGA 708

Qy 239 GTGCAGTGGCGGATTCGACTCAGCGCAACCTCCGCTCCGCGCTTAAGCGATTCTCTCT 298
Db 707 GTGCAGTGGCGGATCTCGGCTCACTGCAAGCTCCGCTCCCGGGTTTCAGGCCATTCTCTCT 648

Qy 299 GCGTCAGCTCCCAAGTAGCTGGGACTACAAGCTCGGGACACACG 344
Db 647 GCGTCAGCTCCCGAGTAGCTGGGACTACAAGCGCCGCTACCAAG 602

RESULT 10
US-09-954-531-146/c
; Sequence 146, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
; APPLICANT: Weaver, Zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Can
; FILE REFERENCE: 689290-77
; CURRENT APPLICATION NUMBER: US/09/954,531
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/60/233,133
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 1392
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 146
; LENGTH: 1160
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-531-146

Query Match          23.9%; Score 119.6; DB 9; Length 1160;
Best Local Similarity 82.5%; Pred. No. 5.6e-24;
Matches 137; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 179 AATTTTGTGTTTGTGTTTGTGTTTGTGAGACAGGTTCTCGAGGTGTCAACCCAGGCTGGA 238
Db 767 ATTATTTTGTGTTTGTGTTTGTGTTTGTGAGACGAGTCTCGCTCTGTGCGCCAGGCTGGA 708

Qy 239 GTGCAGTGGCGGATTCGACTCAGCGCAACCTCCGCTCCGCGCTTAAGCGATTCTCTCT 298
Db 707 GTGCAGTGGCGGATCTCGGCTCACTGCAAGCTCCGCTCCCGGGTTTCAGGCCATTCTCTCT 648
```


RESULT 14

Query Match 23.6%; Score 118.2; DB 10; Length 5159;
Best Local Similarity 82.8%; Pred. NO. 3.2e-23;
Matches 135; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

RESULT 15

Query Match 23.6%; Score 118.2; DB 9; Length 6834;
Best Local Similarity 83.4%; Pred. No. 3.8e-23;
Matches 146; Conservative 0; Mismatches 28; Indels 1; Gaps 1;

[illegible]

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 15:56:54 ; Search time 28.1345 Seconds
(without alignments)

5450.193 Million cell updates/sec

Title: US-09-434-382-28_COPY_1_500

Perfect score: 500

Sequence: 1 taccagtgactgaattcta.....aagctctgagggaactgacgt 500

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	500	100.0	26664	4	US-09-564-805-28
2	120.4	24.1	7680	4	US-09-210-748A-3
3	119.6	23.9	70000	4	US-09-851-896-3
4	115.6	23.1	84495	4	US-09-797-906-3
5	115.4	23.1	8285	4	US-09-732-025-3
6	115.4	23.1	11827	4	US-09-739-455-3
7	115	23.0	38564	4	US-09-734-673-3
8	112.8	22.6	70000	4	US-09-851-896-3
9	111.6	22.3	19011	1	US-08-310-356-36
10	111.6	22.3	19557	5	PCT-US92-06300-1
11	111.4	22.3	2099	4	US-08-938-669A-5
12	111.4	22.3	14636	4	US-09-173-914-6
13	111.2	22.2	246240	2	US-08-724-394A-20
14	111.2	22.2	246240	2	US-08-724-394A-21
15	111.2	22.2	246240	2	US-08-724-394A-22
16	110.8	22.1	5375	3	US-08-757-223-7
17	110.4	22.1	12394	4	US-09-488-856A-10
18	110.4	22.1	84495	4	US-09-797-906-3
19	110.2	22.0	320	1	US-08-629-939-5
20	110.2	22.0	320	1	US-08-759-873-5
21	110	22.0	36159	4	US-09-749-588-3
22	109.8	22.0	246240	2	US-08-724-394A-20
23	109.8	22.0	246240	2	US-08-724-394A-21
24	109.8	22.0	246240	2	US-08-724-394A-22
25	109.4	21.9	568	4	US-09-347-114A-93
26	109.4	21.9	9734	4	US-09-347-114A-80
27	109.2	21.8	14796	4	US-08-975-080-35

c 28	109.2	21.8	14796	4	US-09-630-706-10	Sequence 10, Appl
c 29	109.2	21.8	14796	4	US-09-496-694B-3	Sequence 3, Appl
c 30	109	21.8	602	4	US-09-078-294-27	Sequence 27, Appl
c 31	108.8	21.8	16063	4	US-09-801-052-3	Sequence 3, Appl
c 32	108.6	21.7	36159	4	US-09-749-588-3	Sequence 3, Appl
c 33	108.4	21.7	1624	4	US-08-430-225A-19	Sequence 19, Appl
c 34	108.4	21.7	162450	4	US-09-345-882-1	Sequence 1, Appl
c 35	108.2	21.6	45716	4	US-08-965-048-5	Sequence 5, Appl
c 36	108.2	21.6	45989	4	US-08-965-048-6	Sequence 6, Appl
c 37	108	21.6	16389	4	US-09-741-154-3	Sequence 3, Appl
c 38	108	21.6	18073	4	US-09-078-294-12	Sequence 12, Appl
c 39	108	21.6	50000	4	US-09-146-053-3	Sequence 3, Appl
c 40	108	21.6	112132	4	US-08-741-150-3	Sequence 3, Appl
c 41	107.8	21.6	1001	4	US-09-641-638-629	Sequence 629, App
c 42	107.8	21.6	1001	4	US-09-641-638-630	Sequence 630, App
c 43	107.8	21.6	3742	1	US-08-694-915-5	Sequence 5, Appl
c 44	107.8	21.6	3748	2	US-08-958-240-1	Sequence 1, Appl
c 45	107.8	21.6	20674	4	US-09-641-638-651	Sequence 651, App

ALIGNMENTS

RESULT 1

US-09-564-805-28
; Sequence 28, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigan, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 26664
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (910)..(13104)
; OTHER INFORMATION: exon 1: 910-1154; exon 2: 1736-1786; exon 3:
; OTHER INFORMATION: 1925-1995; exon 4: 3025-3089; exon 5: 4361-4418;
; OTHER INFORMATION: exon 6: 5582-5650; exon 7: 7075-7194; exon 8:
; OTHER INFORMATION: 8186-8244; exon 9: 12878-12936; exon 10:
; NAME/KEY: misc feature
; LOCATION: (13756)..(22917)
; OTHER INFORMATION: exon 11: 13756-13868; exon 12: 15283-15378; exon
; OTHER INFORMATION: 13: 16278-16416; exon 14: 16498-16583; exon 15:
; OTHER INFORMATION: 19583-18701; exon 16: 20349-20445; exon 17:
; OTHER INFORMATION: 22172-22310; exon 18: 22879-22917
; NAME/KEY: misc feature
; LOCATION: (23045)..(26452)
; OTHER INFORMATION: exon 19: 23045-23154; exon 20: 23795-23895; exon
; OTHER INFORMATION: 21: 23973-24093; exon 22: 24354-24432; exon 23:
; OTHER INFORMATION: 25026-25170; exon 24: 25812-26036; polyadenylation
; OTHER INFORMATION: signal: 26447-26452
; NAME/KEY: variation
; LOCATION: (826)..(23879)
; OTHER INFORMATION: s at positions 826 and 23180 is G or C; Y at
; OTHER INFORMATION: positions 1914, 5568, 7165, 16431, 1857 and 20486
; OTHER INFORMATION: is C or T; n at position 13128 is t or tgat; r at
; OTHER INFORMATION: positions 22211 and 23879 is A or G.

US-09-797-906-3

Query Match 23.1%; Score 115.6; DB 4; Length 84495;
Best Local Similarity 66.7%; Pred. No. 4.9e-23;
Matches 180; Conservative 0; Mismatches 89; Indels 1; Gaps 1;
QY 86 TGAGAGTCAAGATCAGGACGATTAACAATGAGCCAGACTCTGCTTTAGACTTGTCC 145
Db 15760 TGGCAAGCTGACCATGACAGATTTGGGAGACACAGCTGATGAAGCATTTTTTCCAA 15819
QY 146 TCTATTCTGCTTTCTTCCAAAAACACTACAATTTTTTGTGTTTGTGTTTGTGTTT 205
Db 15820 TTGTAAATATATTATAACAAGAACATTATATCTTTTCTTTTCTTTTCTTTT 15879
QY 206 GAGACAGGGTCTCAGGTGTCAACCCAGGCTGGAGTGGCGGCGATTCGACTCACCG 265
Db 15880 GAGACAGAGTCTTGCTCTGTCAACCCAGGCTGGAGTGGCGGTGATCTTGGCTCAGGG 15939
QY 266 CAACCTCCGCTCC-GGGCTTAAGCGATTCTCTGCTCAGCTCAGCTCCCAAGTAGCTGGAC 324
Db 15940 CAACCTCCGCTCTCTGGTTTCAAGCGATTCTCTGCTCAGCTCAGCTCCGAGTAGCTGGAC 15999
QY 325 TACAAGCTCGGACACACCACTAAAAATGAT 354
Db 16000 TGCAGGTGGCACCACCATGCCAGCTAAT 16029

RESULT 5

US-09-732-025-3
; Sequence 3, Application US/09732025
; Patent No. 6416990
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001011
; CURRENT APPLICATION NUMBER: US/09/732,025
; CURRENT FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 8285
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(8285)
; OTHER INFORMATION: n = A,T,C or G
US-09-732-025-3

Query Match 23.1%; Score 115.4; DB 4; Length 8285;
Best Local Similarity 74.1%; Pred. No. 2.2e-23;
Matches 146; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
QY 147 CTATTCTGCTTTCTTCCAAAAACACTACAATTTTTTGTGTTTGTGTTTGTGTTT 206
Db 278 CTTTTCTCTTTTGTGAATCTTTTGGACAAAGTAGTTTGTGTTGTTGTTGTTGTT 337
QY 207 AGACAGGGTCTCAGGTGTCAACCCAGGCTGGAGTGGCGGCGATTTGACTCACCGC 266
Db 338 AGACAGGGTCTGCTCTGTCAACCCAGGCTGGAGTGGCGGCGCAATCCAGGCTCACTGC 397
QY 267 AACCTCCGCTCCGCTTAAGCGATTCTCTGCTCAGCTCCCAAGTAGCTGGGACTA 326
Db 398 AACCTCTGCTCCGCTCAAGCGATTCTCTACCTCAGCTCCCAAGTAGCTGGGACAA 457
QY 327 CAAGCTCGGACACCAAC 343
Db 458 CAGGCTCATGCACCAC 474

RESULT 6

QY 127 CTGCTTTAGAGCTTGCTCTATTCTTCTTTTCTTCCAAAAACACTACAATTTT 186

US-09-739-455-3

; Sequence 3, Application US/09739455
; Patent No. 6413756
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL000653
; CURRENT APPLICATION NUMBER: US/09/739,455
; CURRENT FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 11827
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(11827)
; OTHER INFORMATION: n = A,T,C or G
US-09-739-455-3

Query Match 23.1%; Score 115.4; DB 4; Length 11827;
Best Local Similarity 74.1%; Pred. No. 2.6e-23;
Matches 146; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
QY 147 CTATTCTGCTTTCTTCCAAAAACACTACAATTTTTTGTGTTTGTGTTTGTGTTT 206
Db 2129 CTTTTCCTCTTTTGTGAATCTTTTGGACAAAGTAGTTGTTGTTGTTGTTGTT 2188
QY 207 AGACAGGGTCTCAGGTGTCAACCCAGGCTGGAGTGGCGGCGATTTGACTCACCGC 266
Db 2189 AGACAGGGTCTGCTCTGTCAACCCAGGCTGGAGTGGCGGCGCAATCCAGGCTCACTGC 2248
QY 267 AACCTCCGCTCCGCTTAAGCGATTCTCTGCTCAGCTCCCAAGTAGCTGGGACTA 326
Db 2249 AACCTCTGCTCCGCTCAAGCGATCTCTTACCTCAGCTCCCAAGTAGCTGGGACAA 2308
QY 327 CAAGCTCGGACACCAAC 343
Db 2309 CAGGCTCATGTACCAC 2325

RESULT 7

US-09-734-673-3
; Sequence 3, Application US/09734673
; Patent No. 6410294
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001020
; CURRENT APPLICATION NUMBER: US/09/734,673
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 38564
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(38564)
; OTHER INFORMATION: n = A,T,C or G
US-09-734-673-3

Query Match 23.0%; Score 115; DB 4; Length 38564;
Best Local Similarity 67.3%; Pred. No. 5.3e-23;
Matches 177; Conservative 0; Mismatches 85; Indels 1; Gaps 1;

QY 127 CTGCTTTAGAGCTTGCTCTATTCTTCTTTTCTTCCAAAAACACTACAATTTT 186


```
FEATURE:
NAME/KEY: intron
LOCATION: 6935..7758
OTHER INFORMATION: /number=5
FEATURE:
NAME/KEY: exon
LOCATION: 7759..7856
OTHER INFORMATION: /number=6
FEATURE:
NAME/KEY: intron
LOCATION: 7857..9443
OTHER INFORMATION: /number=6
FEATURE:
NAME/KEY: exon
LOCATION: 9444..9573
OTHER INFORMATION: /number=7
FEATURE:
NAME/KEY: intron
LOCATION: 9574..10866
OTHER INFORMATION: /number=7
FEATURE:
NAME/KEY: exon
LOCATION: 10867..11081
OTHER INFORMATION: /number=8
FEATURE:
NAME/KEY: intron
LOCATION: 11082..12480
OTHER INFORMATION: /number=8
FEATURE:
NAME/KEY: exon
LOCATION: 12481..12613
OTHER INFORMATION: /number=9
FEATURE:
NAME/KEY: intron
LOCATION: 12614..13701
OTHER INFORMATION: /number=9
FEATURE:
NAME/KEY: exon
LOCATION: 13702..13799
OTHER INFORMATION: /number=10
FEATURE:
NAME/KEY: intron
LOCATION: 13800..14976
OTHER INFORMATION: /number=10
FEATURE:
NAME/KEY: exon
LOCATION: 14977..15115
OTHER INFORMATION: /number=11
FEATURE:
NAME/KEY: intron
LOCATION: 15116..15533
OTHER INFORMATION: /number=11
FEATURE:
NAME/KEY: exon
LOCATION: 15534..15757
OTHER INFORMATION: /number=12
FEATURE:
NAME/KEY: intron
LOCATION: 15758..16949
OTHER INFORMATION: /number=12
FEATURE:
NAME/KEY: exon
LOCATION: 16950..17082
OTHER INFORMATION: /number=13
FEATURE:
NAME/KEY: intron
LOCATION: 17083..17696
OTHER INFORMATION: /number=13
FEATURE:
NAME/KEY: exon
LOCATION: 17697..17764
OTHER INFORMATION: /number=14
```

```
NAME/KEY: intron
LOCATION: 17765..18534
OTHER INFORMATION: /number=14
FEATURE:
NAME/KEY: CDS
LOCATION: join(1776..1854, 2564..2621, 4076..4208,
LOCATION: 6041..6252, 6802..6934, 7759..7856, 9444..9573,
LOCATION: 10867..11081, 12481..12613, 13702..13799,
LOCATION: 14977..15115, 15534..15757, 16950..17082,
LOCATION: 17697..17741)
OTHER INFORMATION: /product="human serum albumin"
FEATURE:
NAME/KEY: exon
LOCATION: 18535..18697
OTHER INFORMATION: /number=15
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 17742..18697
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1737..1775
PUBLICATION INFORMATION:
AUTHORS: Minichetti, P P
AUTHORS: Ruffner, D E
AUTHORS: Kuang, W-J
AUTHORS: Dennison, O E
AUTHORS: Hawkins, J W
AUTHORS: Beattie, W G
AUTHORS: Dugaiczky, A
TITLE: MOLECULAR STRUCTURE OF THE HUMAN ALBUMIN
TITLE: GENE IS REVEALED BY NUCLEOTIDE SEQUENCE WITHIN
TITLE: Q11-22 OF CHROMOSOME 4
JOURNAL: J. Biol. Chem.
VOLUME: 261
PAGES: 6747-6757
DATE: 1986
RELEVANT RESIDUES IN SEQ ID NO: 36: FROM 1 TO 19011
US-08-310-356-36

Query Match      22.3%; Score 111.6; DB 1; Length 19011;
Best Local Similarity 81.6%; Pred. No. 3.7e-22;
Matches 129; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 187 TTTTGTGTTTGTGTTTGTGACACAGGGTCTCGAGGTGTACCCAGGCTGGAGTGCAGTG 246
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3038 TTAATTTTTTTTTTTTTTAAAGACAGGGTCTCGCTCTGTCGCCAGGCTGGAGTGCAGTG 3097

QY 247 GCGGATTTCGACTCACCAGCAACCTCCGCTCCGGCTTAAGCGATTCTCTGCCTCAGC 306
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3098 GCGCAATCTCGGCTCACTGCAAACTCCGTCGCCGGTTTACGCCATTCTCTGCCTCAGC 3157

QY 307 CTCCCAAGTAGCTGGGACTACAAGCTCGGGACACCACG 344
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3158 CTCCCGAGTAGCTGGGACTACAGCGCCCGCCATCACG 3195

RESULT 10
PCT-US92-06300-1
; Sequence 1, Application PC/TUS9206300
; GENERAL INFORMATION:
; APPLICANT: Hurwitz, David R
; APPLICANT: Nathan, Margret
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: Transgenic Protein Production
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer, Inc.
; STREET: 500 Virginia Ave., Bldg. 3A
; CITY: Ft. Washington
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19034
```


GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 15:56:54 ; Search time 2075.34 Seconds
(without alignments)
11232.524 Million cell updates/sec

Title: US-09-434-382-28_COPY_21800_22600

Perfect score: 801

Sequence: 1 agtgcctgtcctgtatttt.....agcggaagcttgcacggat 801

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rtd.*

36: em.htg.mam.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	800.6	100.0	118788	9	AC005277	AC005277 Homo sapi
2	728.6	91.0	740	9	AF304371S2	AF304369 Homo sapi
3	434.8	54.3	2793	9	AK094012	AK094012 Homo sapi
4	145.8	18.2	2163	9	AK074244	AK074244 Homo sapi
5	145.8	18.2	2734	9	AK094333	AK094333 Homo sapi
6	145.8	18.2	2893	9	AF308694	AF308694 Gorilla g
7	145.8	18.2	2908	9	AF304370	AF304370 Homo sapi
8	145.8	18.2	2908	9	AF308698	AF308698 Pan trogl
9	145.8	18.2	2976	9	AK001392	AK001392 Homo sapi
10	145.8	18.2	2992	6	AX405825	AX405825 Sequence
11	145.8	18.2	2997	9	BC001939	BC001939 Homo sapi
12	145.8	18.2	3006	9	BC004158	BC004158 Homo sapi
13	144.2	18.0	2907	9	AK094687	AK094687 Homo sapi
14	139	17.4	167336	2	AC128361	AC128361 Rattus no
15	138.6	17.3	34593	10	AF348157	AF348157 Mus muscu
16	138.6	17.3	196566	10	AL663045	AL663045 Mouse DNA
17	137	17.1	188836	2	AC097911	AC097911 Rattus no
18	102.2	12.8	2712	10	AF308696	AF308696 Mus muscu
19	46	5.7	125020	9	AF429315	AF429315 Homo sapi
20	45.4	5.7	144048	9	AC013762	AC013762 Homo sapi
21	45	5.6	173016	2	AC019123	AC019123 Homo sapi
22	44.2	5.5	214025	9	AC026782	AC026782 Homo sapi
23	43.8	5.5	120703	9	AC002364	AC002364 Homo sapi
24	43.8	5.5	164846	2	AC023007	AC023007 Homo sapi
25	43.8	5.5	248366	2	AL805901	AL805901 Mus muscu
26	43.4	5.4	179838	2	AC130575	AC130575 Rattus no
27	43	5.4	199847	2	AC127320	AC127320 Mus muscu
28	42.8	5.3	157173	9	AC009501	AC009501 Homo sapi
29	42.6	5.3	182411	9	AC068044	AC068044 Homo sapi
30	42.4	5.3	119191	9	AP004782	AP004782 Homo sapi
31	42.2	5.3	120512	9	AL359852	AL359852 Human DNA
32	42	5.2	183540	2	AC026114	AC026114 Homo sapi
33	42	5.2	188972	9	AC020611	AC020611 Homo sapi
34	41.8	5.2	67970	3	PFMAL1P3	AL031746 Plasmodiu
35	41.8	5.2	138915	2	AF005246	AF005246 Oryza sat
36	41.8	5.2	167764	2	AF004303	AF004303 Oryza sat
37	41.8	5.2	236120	14	AF063866	AF063866 Melanoplu
38	41.6	5.2	161355	2	AC011154	AC011154 Homo sapi
39	41.6	5.2	182374	2	AP003968	AP003968 Homo sapi
40	41	5.1	140106	9	AC096713	AC096713 Homo sapi
41	41	5.1	209060	2	AC019005	AC019005 Homo sapi
42	41	5.1	224448	2	PFMAL4P4	AL035477 Plasmodiu
43	40.8	5.1	206915	10	AL663081	AL663081 Mouse DNA
44	40.8	5.1	210185	2	AC125321	AC125321 Mus muscu
45	40.6	5.1	6370	6	AX344345	AX344345 Sequence

ALIGNMENTS

RESULT 1
AC005277
LOCUS AC005277 118788 bp DNA linear PRI 23-JUL-1998
DEFINITION Homo sapiens chromosome 17, clone hRPK.597_M_12, complete sequence.
ACCESSION AC005277
VERSION AC005277.1 GI:3337311
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 118788)
AUTHORS Birren,B., Fasman,K., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 17, clone hRPK.597_M_12
JOURNAL Unpublished


```

repeat_region complement(21101..21338)
repeat_region /rpt_family="AluSx"
repeat_region complement(21348..21598)
repeat_region /rpt_family="L1M4"
repeat_region 21799..22104
repeat_region /rpt_family="AluJo"
repeat_region 22151..22562
repeat_region /rpt_family="MSTA"
repeat_region 23978..24287
repeat_region /rpt_family="AluSq"
repeat_region 24852..25199
repeat_region /rpt_family="L2"
repeat_region complement(25697..25985)
repeat_region /rpt_family="MUTLD"
unsure 25816..25895
repeat_region /note="Single-stranded coverage."
repeat_region complement(25994..26099)
repeat_region /rpt_family="purine-rich"
repeat_region complement(26122..26291)
repeat_region /rpt_family="MUTLD"
repeat_region 26312..26337
repeat_region /rpt_family="CAAAA)n"
repeat_region 26456..26480
repeat_region /rpt_family="GGGAA)n"
repeat_region complement(26576..26781)
repeat_region /rpt_family="MER6B"
repeat_region complement(26912..26990)
repeat_region /rpt_family="MER21B"
repeat_region 26991..27290
repeat_region /rpt_family="AluSx"
repeat_region complement(27291..27975)
repeat_region /rpt_family="MER21B"
repeat_region 28034..28197
repeat_region /rpt_family="LTR28"
repeat_region 28266..28795
repeat_region /rpt_family="LTR28"
repeat_region 28843..28916
repeat_region /rpt_family="LTR28"
repeat_region 28920..29005
repeat_region /rpt_family="MIR"
repeat_region 29140..29233

Query Match 100.0%; Score 800.6; DB 9; Length 118788;
Best Local Similarity 99.9%; Pred. No. 1.3e-196;
Matches 800; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGTGCCCTGCTGGTATTTTCAACAGAGCTGTGGCCACAGTCAATCTCGATGGTCAGAT 60
Db 99452 AGTGCCCTGCTGGTATTTTCAACAGAGCTGTGGCCACAGTCAATCTCGATGGTCAGAT 99511

QY 61 TCATTGTTAGGACTAATGCTTTAAGCCTCCTATAAACTTTTTTTTTTTTTTTTGTATGC 120
Db 99512 TCATTGTTAGGACTAATGCTTTAAGCCTCCTATAAACTTTTTTTTTTTTTTTTGTATGC 99571

QY 121 CGAGCCTTTGTGTAAGTCTACTTGAAGGGTTTCAGGGTTCCATGATATCTTTTGCTA 180
Db 99572 CGAGCCTTTGTGTAAGTCTACTTGAAGGGTTTCAGGGTTCCATGATATCTTTTGCTA 99631

QY 181 TAAAGAGGATGACATGTAAATACCTTTATGGTTAAATTAATTTGGCTTTATATTAG 240
Db 99632 TAAAGAGGATGACATGTAAATACCTTTATGGTTAAATTAATTTGGCTTTATATTAG 99691

QY 241 CTCCTCAAAAGCAAGCAGGAGACAGAAATTTCTGAGTTGCTTCTGCTCTGTCCTCAA 300
Db 99692 CTCCTCAAAAGCAAGCAGGAGACAGAAATTTCTGAGTTGCTTCTGCTCTGTCCTCAA 99751

QY 301 AGCAGACATCAGCCTCTGAACCATCAGCAGTCTTCTAGTGGCAGTGAATCTTCTCTCT 360
Db 99752 AGCAGACATCAGCCTCTGAACCATCAGCAGTCTTCTAGTGGCAGTGAATCTTCTCTCT 99811

QY 361 TCTCTTCTGAGCCCGCCGACGCTCTGCTACTGGACTGTGGTGGGGACACRTTTGGGCA 420
Db 99812 TCTCTTCTGAGCCCGCCGACGCTCTGCTACTGGACTGTGGTGGGGACACRTTTGGGCA 99871

```

```

QY 421 GCTGTGCCGTCAATTACGAGAGCCAGGTGACAGGGTCTCTGGGCAACCCTGGCTGCTGTGT 480
Db 99872 GCTGTGCCGTCAATTACGAGAGCCAGGTGACAGGGTCTCTGGGCAACCCTGGCTGCTGTGT 99931

QY 481 TGTGTCCCACTGTCACGAGATCACCACGCGTGTGGCTGGGCTGGACCAAAAGCTGG 540
Db 99932 TGTGTCCCACTGTCACGAGATCACCACGCGTGTGGCTGGGCTGGACCAAAAGCTGG 99991

QY 541 AGCCTGGAGGAGGCACTGCCAGTTGAGTTGGCCCTTTGGCTGGCTCTTTTCTCCGCTT 600
Db 99992 AGCCTGGAGGAGGCACTGCCAGTTGAGTTGGCCCTTTGGCTGGCTCTTTTCTCCGCTT 100051

QY 601 CCAAACTTGCACAGAGCTTTTGTACTCATCTCTGGCTAGGAAATGTTTTTTTGAAAAAC 660
Db 100052 CCAAACTTGCACAGAGCTTTTGTACTCATCTCTGGCTAGGAAATGTTTTTTTGAAAAAC 100111

QY 661 TCAACATAGTCTCTGCGCCACAGAAATGCTCTCTCTGTTTCTGTTTCTTCTCTCTG 720
Db 100112 TCAACATAGTCTCTGCGCCACAGAAATGCTCTCTCTGTTTCTGTTTCTTCTCTCTG 100171

QY 721 AGCAGGACAGGTTTGAGTTTACCCAGCCTTCTTCTGAGTCTTGAATCTCACAGGCTGCT 780
Db 100172 AGCAGGACAGGTTTGAGTTTACCCAGCCTTCTTCTGAGTCTTGAATCTCACAGGCTGCT 100231

QY 781 CAGCGAAGCTTTGACCGGAT 801
Db 100232 CAGCGAAGCTTTGACCGGAT 100252

RESULT 2
AF304371S2 740 bp DNA linear PRI 19-OCT-2000
LOCUS Homo sapiens putative prostate cancer susceptibility protein
DEFINITION HPC2/BLAC2 gene, exon 17 and partial cds.
ACCESSION AF304369
VERSION AF304369.1 GI:10880929
KEYWORDS 2 of 2
SEGMENT Homo sapiens.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 740)
AUTHORS Tavtigian,S.V., Simard,J., Teng,D.H.F., Baumgard,M., Beck,A., Camp,N.J., Carillo,A.R., Chen,Y., Dayananth,P., Desrochers,M., Dumont,M., Farnham,J.M., Frank,D., Frye,C., Ghaffari,S., Gupta,J.S., Hu,R., Iliev,D., Janacki,T., Kort,E.N., Laity,K.E., Leavitt,A., Leblanc,G., McArthur-Morrison,J., Pederson,A., Penn,B., Peterson,K.T., Reid,J.E., Richards,S., Schroeder,M., Smith,R., Snyder,S.C., Swedlund,B., Swensen,J., Thomas,A., Tranchant,M., Woodland,A.-M., Labrie,F., Skolnick,M.H., Neuhausen,S., Rommens,J. and Cannon-Albright,L.A.
TITLE A strong candidate prostate cancer susceptibility gene at chromosome 17p
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 740)
AUTHORS Tavtigian,S.V., Simard,J., Teng,D.H.F., Baumgard,M., Beck,A., Camp,N.J., Carillo,A.R., Chen,Y., Dayananth,P., Desrochers,M., Dumont,M., Farnham,J.M., Frank,D., Frye,C., Ghaffari,S., Gupta,J.S., Hu,R., Iliev,D., Janacki,T., Kort,E.N., Laity,K.E., Leavitt,A., Leblanc,G., McArthur-Morrison,J., Pederson,A., Penn,B., Peterson,K.T., Reid,J.E., Richards,S., Schroeder,M., Smith,R., Snyder,S.C., Swedlund,B., Swensen,J., Thomas,A., Tranchant,M., Woodland,A.-M., Labrie,F., Skolnick,M.H., Neuhausen,S., Rommens,J. and Cannon-Albright,L.A.
TITLE Direct Submission
JOURNAL Submitted (12-SEP-2000) Myriad Genetics, Inc., 320 Wakara Way, Salt Lake City, UT 84108, USA
FEATURES
source Location/Qualifiers
1..740
/organism="Homo sapiens"
/db_xref="taxon:9606"

```


Qy	458	CTGGGCAACCCCTGGCTGCTGTTGTTGTCTCCACCTGCACGCAGATCAACACGG	512
Db	845	CTGGGCAACCCCTGGCTGCTGTTGTTGTCTCCACCTGCACGCAGATCAACACGG	899
RESULT 5			
AK094333			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
SOURCE			
BASE COUNT			
ORIGIN			
Query Match			
Best Local Similarity			
Matches			
Qy	338	AGTGGCAGTGA	397
Db	1314	ATTGAAATGTCA	1373
Qy	398	TGTGTTGAGGCAC	457
Db	1374	TGTGTTGAGGCAC	1433
Qy	458	CTGGGCAACCCCTGGCTGCTGTTGTTGTCTCCACCTGCACGCAGATCAACACGG	512
Db	1434	CTGGGCAACCCCTGGCTGCTGTTGTTGTCTCCACCTGCACGCAGATCAACACGG	1488
RESULT 6			

[illegible]

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GYER
Center clone name: CH230-76E21
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 109154 bases at least Q40
Consensus quality: 116608 bases at least Q30
Consensus quality: 121720 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
NOTE: This is a 'working draft' sequence. It currently consists of 72 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

*	1	1504:	contig	of 1504 bp in length
*	1505	1604:	gap of unknown length	
*	1605	2700:	contig of 1036 bp in length	
*	2701	2800:	gap of unknown length	
*	2801	4303:	contig of 1509 bp in length	
*	4310	4403:	gap of unknown length	
*	4410	5676:	contig of 1267 bp in length	
*	5677	5776:	gap of unknown length	
*	5777	7510:	contig of 1734 bp in length	
*	7511	7610:	gap of unknown length	
*	7611	8958:	contig of 1348 bp in length	
*	8959	9058:	gap of unknown length	
*	9059	10177:	contig of 1119 bp in length	
*	10178	10277:	gap of unknown length	
*	10278	11696:	contig of 1419 bp in length	
*	11697	11796:	gap of unknown length	
*	11797	13134:	contig of 1338 bp in length	
*	13135	13234:	gap of unknown length	
*	13235	14464:	contig of 1230 bp in length	
*	14465	14564:	gap of unknown length	
*	14565	15674:	contig of 1110 bp in length	
*	15675	15774:	gap of unknown length	
*	15775	16783:	contig of 1009 bp in length	
*	16784	16883:	gap of unknown length	
*	16884	18010:	contig of 1127 bp in length	
*	18011	18110:	gap of unknown length	
*	18111	19878:	contig of 1768 bp in length	
*	19879	19978:	gap of unknown length	
*	19979	21592:	contig of 1614 bp in length	
*	21593	21692:	gap of unknown length	
*	21693	22664:	contig of 1172 bp in length	
*	22665	22964:	gap of unknown length	
*	22965	25063:	contig of 2099 bp in length	
*	25064	25163:	gap of unknown length	
*	25164	26191:	contig of 1028 bp in length	
*	26192	26291:	gap of unknown length	
*	26292	27604:	contig of 1313 bp in length	
*	27605	27704:	gap of unknown length	
*	27705	29172:	contig of 1468 bp in length	
*	29173	29272:	gap of unknown length	
*	29273	30658:	contig of 1386 bp in length	
*	30659	30758:	gap of unknown length	
*	30759	31911:	contig of 1153 bp in length	
*	31912	32011:	gap of unknown length	
*	32012	33192:	contig of 1181 bp in length	
*	33193	33292:	gap of unknown length	
*	33293	34521:	contig of 1229 bp in length	
*	34522	34621:	gap of unknown length	

*	34622	35930: contig of 1309 bp in length
*	35931	36030: gap of unknown length
*	36031	37423: contig of 1393 bp in length
*	36032	37423: contig of 1393 bp in length
*	37424	37523: gap of unknown length
*	37524	38838: contig of 1315 bp in length
*	38839	38938: gap of unknown length
*	38939	38938: gap of unknown length
*	40679	40678: contig of 1740 bp in length
*	40779	40778: gap of unknown length
*	42403	42403: contig of 1625 bp in length
*	42404	42503: gap of unknown length
*	42504	44219: contig of 1716 bp in length
*	44220	44319: gap of unknown length
*	44320	47210: contig of 2891 bp in length
*	47321	47310: gap of unknown length
*	47311	49688: contig of 2378 bp in length
*	49689	49788: gap of unknown length
*	49789	51816: contig of 2028 bp in length
*	51817	51916: gap of unknown length
*	51917	53775: contig of 1859 bp in length
*	53776	53875: gap of unknown length
*	53876	55631: contig of 1756 bp in length
*	55632	55731: gap of unknown length
*	55732	57665: contig of 1934 bp in length
*	57666	57765: gap of unknown length
*	57766	60037: contig of 2362 bp in length
*	60028	60137: gap of unknown length
*	60128	62257: contig of 2130 bp in length
*	62258	62357: gap of unknown length
*	62358	64599: contig of 2202 bp in length
*	64560	64659: gap of unknown length
*	64660	66467: contig of 1808 bp in length
*	66568	66567: gap of unknown length
*	66568	68164: contig of 1597 bp in length
*	68165	68264: gap of unknown length
*	88265	89323: contig of 1059 bp in length
*	89324	89423: gap of unknown length
*	89424	71645: contig of 2222 bp in length
*	71646	71745: gap of unknown length
*	71746	74023: contig of 2278 bp in length
*	74024	74123: gap of unknown length
*	74124	76416: contig of 2293 bp in length
*	76417	76516: gap of unknown length
*	76517	79082: contig of 2366 bp in length
*	79083	79182: gap of unknown length
*	79183	82081: contig of 2899 bp in length
*	82082	82181: gap of unknown length
*	82182	84830: contig of 2639 bp in length
*	84821	84920: gap of unknown length
*	84921	87603: contig of 2883 bp in length
*	87604	87703: gap of unknown length
*	87704	89743: contig of 2040 bp in length
*	89744	89843: gap of unknown length
*	89844	92863: contig of 3020 bp in length
*	92864	92963: gap of unknown length
*	92964	95013: contig of 2050 bp in length
*	95014	95113: gap of unknown length
*	95114	97602: contig of 2489 bp in length
*	97603	97702: gap of unknown length
*	97703	100852: contig of 3160 bp in length
*	100863	100962: gap of unknown length
*	100963	102803: contig of 1841 bp in length
*	102804	102903: gap of unknown length
*	102904	105239: contig of 2396 bp in length
*	105300	105399: gap of unknown length
*	105400	108267: contig of 2868 bp in length
*	108268	108367: gap of unknown length

Query Match 17.4%; Score 139; DB 2; Length 167336;
Best Local Similarity 62.5%; Pred. No. 5.1e-25;
Matches 267; Conservative 0; Mismatches 155; Indels 5;

Qy 99 TTTT TTTT TTTT TTTT TGATGCCAGCCTT - TGTGTAAGTCTACTTGAAAGGGTTTCAGG 157

```
Db 127052 TTCCTTCTCTGCTTTGATGCCAGCATGTGTTGATGCATATCCTTAAAGGCTTCAGG 126993
Oy 158 GTTCCATGATGACTCTTTGCTATTAAGAGGATGACACATGTAAATACACCTTTATGGTT 217
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 126992 GTTCTTTGGATTTTCCCTTAAACAGAGAAGACAGACTGCTCTTAAACGACTCATGTGGTT 126933
Oy 218 AAATTAATGGCTTTTATATTAGCTCCTCAAGCAAGCAGGAGACAGAAATTTCTGC 277
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 126932 AAG-TAGTTGACTTCATATGCACTCCACAAAAGATGGTGGGTG---AGCTGTGGAA 126877
Oy 278 AGTTGCTTCTTGCTGCTGTGTCACCAAGCAGACATCAGGCTCTGAACCATCAGCAGTCTTCT 337
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 126876 GATGACGTGGTGCCACAGATTGCTGCTGCTGCTGCTGCTGCAACATGCCCTTCTG 126817
Oy 338 AGTGGCAGTACTCTTCTCTCTTCTGTCAGCCCGCAGCAGCTGCTGCTGCTGCTGAC 397
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 126816 GGAAGTCTAAAGCCCTGTTGTGACTGCTTCTGACGCCCTGACAAGTCTGTGCTGTGAT 126757
Oy 398 TGTGTGAGGGCACRTTTGGCGAGCTGTGCCCTATTAGCGAGACCGGTGCACAGGGTC 457
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 126756 TGTGAGAGAGGCACCTTTTGGCGAGTTATGCCGTCTATTGAGCAGCAATACAGAGATC 126697
Oy 458 CTGGCACCCTCGCTGCTGTGTTGTGCTCCACCTGCAGCAGATCACCACACGGTGAAT 517
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 126696 CTCTGCACTTACTGCTGTGTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 126637
Oy 518 GTTGGGC 524
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 126636 GTGAGAC 126630

RESULT 15
AF348157
LOCUS
DEFINITION
Mus musculus putative prostate cancer susceptibility protein
(Ela2) gene, complete cds, alternatively spliced.
ACCESSION
AF348157
VERSION
AF348157.1 GI:13540341
KEYWORDS
Mus musculus.
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (sites)
Frank,D.C., Swedlund,B., Dumont,M., Tavtigian,S.V., Simard,J.,
Teng,D.H.F., Baumgard,M., Beck,A., Camp,N.J., Carillo,A.R.,
Chen,Y., Dayananth,P., Desrochers,M., Farnham,J.M., Frye,C.,
Gaffari,S., Gupta,J.S., Hu,R., Iliev,D., Janecki,T., Kort,E.N.,
Laity,K.E., Leavitt,A., Leblanc,G., McArthur-Morrison,J.,
Pederson,A., Penn,B., Peterson,K.T., Reid,J.E., Richards,S.,
Schroeder,M., Smith,R., Snyder,S.C., Swensen,J., Thomas,A.,
Tranchant,M., Woodland,A.-M., Labrie,F., Skolnick,M.H.,
Neuhausen,S., Rommens,J. and Cannon-Albright,L.A.
Mouse Ela2-containing genomic DNA
unpublished
TITLE
2 (bases 1 to 34593)
JOURNAL
Frank,D.C., Swedlund,B., Dumont,M., Tavtigian,S.V., Simard,J.,
Teng,D.H.F., Baumgard,M., Beck,A., Camp,N.J., Carillo,A.R.,
Chen,Y., Dayananth,P., Desrochers,M., Farnham,J.M., Frye,C.,
Gaffari,S., Gupta,J.S., Hu,R., Iliev,D., Janecki,T., Kort,E.N.,
Laity,K.E., Leavitt,A., Leblanc,G., McArthur-Morrison,J.,
Pederson,A., Penn,B., Peterson,K.T., Reid,J.E., Richards,S.,
Schroeder,M., Smith,R., Snyder,S.C., Swensen,J., Thomas,A.,
Tranchant,M., Woodland,A.-M., Labrie,F., Skolnick,M.H.,
Neuhausen,S., Rommens,J. and Cannon-Albright,L.A.
Direct Submission
TITLE
Submitted (09-FEB-2001) Myriad Genetics, Inc., 320 Wakara Way, Salt
Lake City, UT 84108, USA
JOURNAL
Location/Qualifiers
1. 34593
FEATURES
source
/organism="Mus musculus"
/db_xref="taxon:10090"
gene
<8470..>31277
```

```
/gene="Ela2"
join(<8470..8690,9077..9127,9266..9336,9948..10012,
10957..11014,12096..12164,12933..13064,14902..14960,
16609..16667,16746..16818,17370..17482,18671..18766,
21539..21677,21757..21842,23553..23671,24619..24715,
27291..27429,27763..27801,27930..28039,28553..28652,
28733..28853,29101..29179,29710..29854,30406..30624,
31242..>31277)
/gene="Ela2"
/product="putative prostate cancer susceptibility protein"
/note="alternatively spliced 25 exon form"
join(8470..8690,9077..9127,9266..9336,9948..10012,
10957..11014,12096..12164,12933..13064,14902..14960,
16609..16667,16746..16818,17370..17482,18671..18766,
21539..21677,21757..21842,23553..23671,24619..24715,
27291..27429,27763..27801,27930..28039,28553..28652,
28733..28853,29101..29179,29710..29854,30406..30624,
31242..31277)
/gene="Ela2"
/note="25 exon form"
/product="putative prostate cancer susceptibility protein"
/protein_id="AAK29420.1"
/db_xref="GI:13540342"
/translation="MWALRSLLRPLGLRTMSQGSARRPPPKDPLRLHRTREKRGPG
GGPNTVYLVVAGGRDAGALVFESEYNYLFNCGEGVORLMOEHLKVARLDNI FL
TRHWSNVGGLCGMILITKETGLPKCVLSGGPQLEKYLEAIKIFSGPLKGLAVRPH
SAPEYKDETMTVYQVPIHSERRCGKQSPRTSPNRLSPKQSDSGSANGQCPE
DSSAGANRKAWGRDPSLVAVCKHLRKNFLVKAKELGLPVGTAAIPIAAVKD
GKSITYEGREIAAEELCTPPDPGLVFIWVECPDEGFLPICENDTFKRYQAEAPVA
LVVHIAPESVLIDRSYQWMEFGDPQTHLILNENCPNVHNRSHKIQTQSLIHPDI
PQLTFSVSKEEGSLVPTVRGECILKYQLRPKEWQDITLDCNTDEFTAEALELP
SFQESVEEYKRVNVOENPAPAKRSQYPEIVFLGTGSAIPMKIRNVSTLVNLSPKSV
LLDCGEGTFOQLCRHYGQIDRVLCSTLTVFVSHLHADHGLLNILLQREHALSLG
KPQPLLVAVPTQLRAWLQYHNHCQIILHHVSMIPAKCLOKGAESVNTLERLISLL
LETCDLEEFQTCVLRHCKHAFGCAVHSSGMKVVYSGDTPCEALVQMGKATLIIHE
ATLEDGLEEAEVEXKTHSTTQAINVGMNMAEFIMLNFHSORYAKIPLFSPDFNEKVG
IAFDHMKVCFGDEPTVPKLIPLKALFAGDIEENVREKRELRLVRAALLTQQADSP
EDREPQOKRAHTDEPHSPQSKKESVANTLGARV"
join(<8470..8690,9077..9127,9266..9336,9948..10012,
10957..11014,12096..12164,12933..13064,14902..14960,
16609..16667,16746..16818,17370..17482,18671..18766,
21539..21677,21757..21842,23553..23671,24619..24715,
27291..27429,27763..27801,27930..28039,28553..28652,
28733..28853,29101..29179,29710..29854,30406..30639)
/gene="Ela2"
/product="putative prostate cancer susceptibility protein"
/note="alternatively spliced 24 exon form"
join(8470..8690,9077..9127,9266..9336,9948..10012,
10957..11014,12096..12164,12933..13064,14902..14960,
16609..16667,16746..16818,17370..17482,18671..18766,
21539..21677,21757..21842,23553..23671,24619..24715,
27291..27429,27763..27801,27930..28039,28553..28652,
28733..28853,29101..29179,29710..29854,30406..30639)
/gene="Ela2"
/product="putative prostate cancer susceptibility protein"
/note="24 exon form"
/codon_start=1
/product="putative prostate cancer susceptibility protein"
/protein_id="AAK29421.1"
/db_xref="GI:13540343"
/translation="MWALRSLLRPLGLRTMSQGSARRPPPKDPLRLHRTREKRGPG
GGPNTVYLVVAGGRDAGALVFESEYNYLFNCGEGVORLMOEHLKVARLDNI FL
TRHWSNVGGLCGMILITKETGLPKCVLSGGPQLEKYLEAIKIFSGPLKGLAVRPH
SAPEYKDETMTVYQVPIHSERRCGKQSPRTSPNRLSPKQSDSGSANGQCPE
DSSAGANRKAWGRDPSLVAVCKHLRKNFLVKAKELGLPVGTAAIPIAAVKD
GKSITYEGREIAAEELCTPPDPGLVFIWVECPDEGFLPICENDTFKRYQAEAPVA
LVVHIAPESVLIDRSYQWMEFGDPQTHLILNENCPNVHNRSHKIQTQSLIHPDI
PQLTFSVSKEEGSLVPTVRGECILKYQLRPKEWQDITLDCNTDEFTAEALELP
SFQESVEEYKRVNVOENPAPAKRSQYPEIVFLGTGSAIPMKIRNVSTLVNLSPKSV
LLDCGEGTFOQLCRHYGQIDRVLCSTLTVFVSHLHADHGLLNILLQREHALSLG
KPQPLLVAVPTQLRAWLQYHNHCQIILHHVSMIPAKCLOKGAESVNTLERLISLL
LETCDLEEFQTCVLRHCKHAFGCAVHSSGMKVVYSGDTPCEALVQMGKATLIIHE
ATLEDGLEEAEVEXKTHSTTQAINVGMNMAEFIMLNFHSORYAKIPLFSPDFNEKVG
```


GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 15:56:54 ; Search time 1295.47 Seconds
(without alignments)
11232.524 Million cell updates/sec

Title: US-09-434-382-28_COPY_1_500

Perfect score: 500

Sequence: 1 taccagtgactgaattctta.....aagctctgaggactgacgt 500

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vl.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rtd.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	500	100.0	118788	9	AC005277	AC005277 Homo sapi
2	131	26.2	157159	2	AC026279	AC026279 Homo sapi
3	131	26.2	157590	9	AC087525	AC087525 Homo sapi
4	131	26.2	171418	2	AC023591	AC023591 Homo sapi
5	128.6	25.7	180707	2	AC017038	AC017038 Homo sapi
6	128.6	25.7	184338	9	AC060764	AC060764 Homo sapi
7	126.4	25.3	110000	2	AC009771	Continuation (3 of
8	126.4	25.3	110000	2	AC012386	Continuation (3 of
9	126.2	25.2	148068	9	CNS01RHN_o	AC012386 Homo sapi
10	126.2	25.2	179937	9	CNS05TCV	AL161871 Human chr
11	125.8	25.2	115396	9	AC010366	AL355840 Human chr
12	125.8	25.2	214139	2	AC022129	AC022129 Homo sapi
13	125.6	25.1	74089	9	AL589762	AL589762 Human DNA
14	125.6	25.1	208309	9	CNS01DSC	AL121767 Human chr
15	125.4	25.1	153185	2	AC073228	AC073228 Homo sapi
16	125.4	25.1	178653	2	AC090088	AC090088 Homo sapi
17	125.4	25.1	178980	9	AP001591	AP001591 Homo sapi
18	125.4	25.1	184057	9	AP000753	AP000753 Homo sapi
19	125.4	25.1	184059	2	AC090382	AC090382 Homo sapi
20	125	25.0	65612	9	AL136304	AL136304 Human DNA
21	124.8	25.0	199882	9	AL354720	AL354720 Human DNA
22	124.6	24.9	191468	9	AC092631	AC092631 Homo sapi
23	124.2	24.9	171779	2	AC016000	AC016000 Homo sapi
24	124.2	24.8	50721	2	AC069106	AC069106 Homo sapi
25	124.2	24.8	96512	9	AC022539	AC022539 Homo sapi
26	124.2	24.8	154359	2	AC021718	AC021718 Homo sapi
27	124.2	24.8	158568	9	AC069541	AC069541 Homo sapi
28	124.2	24.8	159895	9	AC091772	AC091772 Homo sapi
29	124.2	24.8	189786	2	AC068015	AC068015 Homo sapi
30	124	24.8	175642	2	AC116934	AC116934 Papio cyn
31	124	24.8	181532	9	AL499582	AL499582 Human DNA
32	123.6	24.7	59915	2	AC111187	AC111187 Homo sapi
33	123.4	24.7	146957	9	AC084264	AC084264 Homo sapi
34	123.4	24.7	186091	9	AC011242	AC011242 Homo sapi
35	123.2	24.6	149679	2	AC080115	AC080115 Homo sapi
36	123.2	24.6	155276	2	AC015576	AC015576 Homo sapi
37	123.2	24.6	161149	2	AC012028	AC012028 Homo sapi
38	123.2	24.6	165434	9	AC022448	AC022448 Homo sapi
39	123.2	24.6	178516	9	AC109993	AC109993 Homo sapi
40	123.2	24.6	190740	2	HSAS24G17	AL457222 Homo sapi
41	123	24.6	164084	9	AL354953	AL354953 Human DNA
42	123	24.6	173853	2	AC018999	AC018999 Homo sapi
43	123	24.6	193803	9	AC022819	AC022819 Homo sapi
44	122.8	24.6	178530	9	AC018738	AC018738 Homo sapi
45	122.8	24.6	188025	2	AC074185	AC074185 Homo sapi

ALIGNMENTS

RESULT 1
AC005277 AC005277 118788 bp DNA linear PRI 23-JUL-1998
LOCUS Homo sapiens chromosome 17, clone hRPK.597_M_12, complete sequence.
DEFINITION AC005277
ACCESSION AC005277
VERSION AC005277.1 GI:3337311
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 118788)
REFERENCE
AUTHORS Birren,B., Fasman,K., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 17, clone hRPK.597_M_12
JOURNAL Unpublished

```

REFERENCE
AUTHORS
2 (bases 1 to 118788)
Birren,B., Fasman,K., Linton,L., Nusbaum,C., Lander,B., Allen,N.,
Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Boatn,C.,
Boutwell,C., Brown,A., Castle,A., Cerny,J., Cooke,P., Depayre,E.,
Devon,K., Dewar,K., Donelan,L., Etemadi,S., Ferreira,P.,
FitzHugh,W., Forrest,C., Funke,R., Gage,D., Gardyna,S.,
Gensheimer,S., Geraghty,K., Gilmartin,T., Grant,G., Hagos,B.,
Harris,K., Horton,L., Howland,J.C., Hui,L., Jacotot,L., Kann,L.,
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
Melldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
Nachman,A., Nahf,R., Naylor,J., Niloff,M., O'Connor,T., Pavlin,B.,
Peterson,K., Riley,R., Roberts,D., Rossello,R., Roy,A., Shyam,R.,
Stange-Thomann,N., Stiwell,J., Stojanovic,N., Stone,C.,
Strickland,C., Subramanian,A., Torruella-Miller,I., Vassiliev,H.,
Vo,A., Wagner,A., Wang,B., Wheeler,J., Wu,Y., Ye,W.J., Zhao,J. and
Zody,M.
TITLE
JOURNAL
Direct Submission
Submitted (10-JUL-1998) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE
AUTHORS
3 (bases 1 to 118788)
Birren,B., Fasman,K., Linton,L., Nusbaum,C., Lander,B., Allen,N.,
Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Boatn,C.,
Boutwell,C., Brown,A., Castle,A., Cerny,J., Cooke,P., Depayre,E.,
Devon,K., Dewar,K., Donelan,L., Etemadi,S., Ferreira,P.,
FitzHugh,W., Forrest,C., Funke,R., Gage,D., Gardyna,S.,
Gensheimer,S., Geraghty,K., Gilmartin,T., Grant,G., Hagos,B.,
Harris,K., Horton,L., Howland,J.C., Hui,L., Jacotot,L., Kann,L.,
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
Melldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
Nachman,A., Nahf,R., Naylor,J., Niloff,M., O'Connor,T., Pavlin,B.,
Peterson,K., Riley,R., Roberts,D., Rossello,R., Roy,A., Shyam,R.,
Stange-Thomann,N., Stiwell,J., Stojanovic,N., Stone,C.,
Strickland,C., Subramanian,A., Torruella-Miller,I., Vassiliev,H.,
Vo,A., Wagner,A., Wang,B., Wheeler,J., Wu,Y., Ye,W.J., Zhao,J. and
Zody,M.
TITLE
JOURNAL
Direct Submission
Submitted (23-JUL-1998) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT
On Jul 23, 1998 this sequence version replaced gi:3335015.
All repeats were identified using RepeatMasker: Smit, A.F.A. &
Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Only the first 118.8 kilobases of this clone are being submitted.
The remainder overlaps accession number AC005274 (WICGR project
L350).

FEATURES
source
1. 118788
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="17"
/clone="hrpk:597 M 12"
/clone.lib="RPC1-11 human BAC library"
795..884
/rpt_family="MIR"
repeat_region
complement(932..1000)
/rpt_family="MIR"
repeat_region
2044..2076
/rpt_family="AT_rich"
repeat_region
2237..2713
/rpt_family="MER34"
repeat_region
2718..3477
/rpt_family="L2"
repeat_region
complement(3478..3821)
/rpt_family="MLT1C"
repeat_region
3822..4007
/rpt_family="L2"
repeat_region
complement(4015..4281)
/rpt_family="MER89"
repeat_region
complement(4282..4321)
/rpt_family="(TAAA)n"
repeat_region
complement(4328..4608)
/rpt_family="L2"
complement(4609..4895)
/rpt_family="MER89"
complement(4896..5186)
/rpt_family="AluSp"
5219..5479
/rpt_family="L2"
complement(5481..5740)
/rpt_family="MER34"
complement(5963..6275)
/rpt_family="AluSp"
complement(6411..6494)
/rpt_family="MER5A"
6830..7013
/rpt_family="MLT1C"
7014..7318
/rpt_family="AluYb8"
7319..7581
/rpt_family="MLT1C"
7721..8032
/rpt_family="AluJb"
8039..8333
/rpt_family="AluSq"
9431..9460
/rpt_family="AT_rich"
9662..9829
/rpt_family="MIR"
complement(11302..11728)
/rpt_family="MER39"
complement(11824..11884)
/rpt_family="LTR16C"
complement(11928..12108)
/rpt_family="LTR16C"
12453..12725
/rpt_family="AluJb"
13377..13785
/rpt_family="WSTB"
14118..14153
/rpt_family="(CA)n"
complement(15320..15620)
/rpt_family="AluJo"
15683..15759
/rpt_family="(CA)n"
complement(15838..15893)
/rpt_family="MER91A"
complement(15915..16080)
/rpt_family="MER91A"
complement(16474..16544)
/rpt_family="L2"
17907..17920
/rpt_family="L1ME"
18271..18311
/rpt_family="L2"
complement(18368..18529)
/rpt_family="AluJb"
complement(18530..18560)
/rpt_family="(CA)n"
complement(18561..18681)
/rpt_family="AluJb"
18760..19065
/rpt_family="AluSg"
complement(19638..19896)
/rpt_family="MIR"
complement(19971..20260)
/rpt_family="L1ME3A"
20263..20552
/rpt_family="MER58B"
complement(20553..20816)
/rpt_family="L1ME3A"
complement(20821..21069)
/rpt_family="L1M4"
complement(21071..21100)
/rpt_family="(TAAA)n"

```

```
repeat_region complement(21101..21338)
repeat_region /rpt_family="AluSx"
repeat_region complement(21348..21598)
repeat_region /rpt_family="L1M4"
repeat_region 21799..22104
repeat_region /rpt_family="AluJo"
repeat_region 22151..22562
repeat_region /rpt_family="MSTA"
repeat_region 23978..24287
repeat_region /rpt_family="AluSq"
repeat_region 24852..25199
repeat_region /rpt_family="L2"
repeat_region complement(25697..25985)
repeat_region /rpt_family="MUT1D"
unsure 25816..25895
repeat_region /note="Single-stranded coverage."
repeat_region complement(25994..26099)
repeat_region /rpt_family="purine-rich"
repeat_region complement(26122..26291)
repeat_region /rpt_family="MUT1D"
repeat_region 26312..26337
repeat_region /rpt_family="(CAAAA)n"
repeat_region 26456..26480
repeat_region /rpt_family="(CGGAA)n"
repeat_region complement(26576..26781)
repeat_region /rpt_family="MER6B"
repeat_region complement(26912..26990)
repeat_region /rpt_family="MER21B"
repeat_region 26991..27290
repeat_region /rpt_family="AluSx"
repeat_region complement(27231..27975)
repeat_region /rpt_family="MER21B"
repeat_region 28034..28197
repeat_region /rpt_family="LTR28"
repeat_region 28266..28795
repeat_region /rpt_family="LTR28"
repeat_region 28843..28916
repeat_region /rpt_family="LTR28"
repeat_region 28920..29005
repeat_region /rpt_family="MIR"
repeat_region 29140..29233

Query Match 100.0%; Score 500; DB 9; Length 118788;
Best Local Similarity 100.0%; Pred. No. 1.le-116;
Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TATCAGGTGACTGAATTCCTATATCTTGAAGTAGGAGATCTGTTATTGCTTTATTACAT 60
Db 77653 TATCAGGTGACTGAATTCCTATATCTTGAAGTAGGAGATCTGTTATTGCTTTATTACAT 77712

QY 61 TTTACATAAGAAAGCTGAGGCTCTGAGAGGTCAAGATCACGACGTAAACAAATGAGCC 120
Db 77713 TTTACATAAGAAAGCTGAGGCTCTGAGAGGTCAAGATCACGACGTAAACAAATGAGCC 77772

QY 121 AGACTCTTGCTTTAGAGCTTGCTCTATCTTGCTTTTCTTTTCCAAAAACACTACAA 180
Db 77773 AGACTCTTGCTTTAGAGCTTGCTCTATCTTGCTTTTCTTTTCCAAAAACACTACAA 77832

QY 181 TTTTCTGTTTGTGTTTGTGTTTTCAGACAGGCTCTCGAGGTGCACCCAGGCTGAGT 240
Db 77833 TTTTCTGTTTGTGTTTGTGTTTTCAGACAGGCTCTCGAGGTGCACCCAGGCTGAGT 77892

QY 241 GCAGTGGCGGATTTGCACTCACCGCAACCTCGCGCTTCGCGCTTAAGGATTCCTCTGC 300
Db 77893 GCAGTGGCGGATTTGCACTCACCGCAACCTCGCGCTTCGCGCTTAAGGATTCCTCTGC 77952

QY 301 CTCAGCTCCCAAGTAGTGGGACTCAAGCTCGGACACACCGTAAAAATGATCAAGTT 360
Db 77953 CTCAGCTCCCAAGTAGTGGGACTCAAGCTCGGACACACCGTAAAAATGATCAAGTT 78012

QY 361 CTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
Db 78013 CTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 78072
```

```
QY 421 CTCAATACAATTGATTCTCTACATTTAATCTCAACCTCAACACCTCAACCTCTTAATCTC 480
Db 78073 CTCAATACAATTGATTCTCTACATTTAATCTCAACCTCAACACCTCAACCTCTTAATCTC 78132

QY 481 AGCTCTGAGGACTGAGCT 500
Db 78133 AGCTCTGAGGACTGAGCT 78152

RESULT 2
AC026279 Homo sapiens chromosome 15 clone RP11-321G12, WORKING DRAFT
LOCUS AC026279
DEFINITION AC026279
ACCESSION AC026279.4 GI:8954310
VERSION AC026279
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 157159)
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 157159)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (21-MAR-2000) Genome Sequencing Center, Washington
MO 63108, USA
COMMENT On Jul 7, 2000 this sequence version replaced gi:7523953.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site:http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH0312G12
----- Summary Statistics -----
Sequencing vector: M13; 86%
Sequencing vector: plasmid; 14%
Chemistry: Dye-primer ET; 86% of reads
Chemistry: Dye-terminator Big Dye; 14% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 148266 bases at least Q40
Consensus quality: 151490 bases at least Q30
Consensus quality: 153444 bases at least Q20
Insert size: 165000; agarose-fp
Insert size: 156059; sum-of-contigs
Quality coverage: 4.15 in Q20 bases; agarose-fp
Quality coverage: 3.96 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1812: contig of 1812 bp in length
* 1813: gap of unknown length
* 1913: contig of 2138 bp in length
* 4050: gap of unknown length
* 4151: gap of unknown length
* 8892: contig of 4742 bp in length
* 8893: gap of unknown length
* 8993: contig of 5346 bp in length
* 14338: gap of unknown length
* 14439: contig of 5622 bp in length
* 20061: gap of unknown length
* 20161: contig of 7084 bp in length
```

```

* 27245 27344: gap of unknown length
* 27345 42412: contig of 15068 bp in length
* 42413 42512: gap of unknown length
* 42513 55808: contig of 13296 bp in length
* 55809 55908: gap of unknown length
* 55909 69598: contig of 13690 bp in length
* 69599 69698: gap of unknown length
* 69699 88164: contig of 18466 bp in length
* 88165 88264: gap of unknown length
* 88265 106938: contig of 18674 bp in length
* 106939 107039: gap of unknown length
* 107039 157159: contig of 50121 bp in length.

```

FEATURES

Source

```

1. 157159
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /chromosome="15"
  /clones="RP11-321G12"
1. 1812
  /note="assembly_name:Contig1"
1913. 4050
  /note="assembly_name:Contig2"
4151. 8892
  /note="assembly_name:Contig3"
8993. 14338
  /note="assembly_name:Contig4"
14439. 20060
  /note="assembly_name:Contig5"
20161. 27244
  /note="assembly_name:Contig6"
27345. 42412
  /note="assembly_name:Contig7"
42513. 55808
  /note="assembly_name:Contig8"
55909. 69598
  /note="assembly_name:Contig9"
69699. 88164
  /note="assembly_name:Contig10"
88265. 106938
  /note="assembly_name:Contig11
clone_end:SP6
vector_side:left"
107039. 157159
  /note="assembly_name:Contig12
clone_end:I7
vector_side:left"

```

```

BASE COUNT 42295 a 34576 c 35453 g 43728 t 1107 others
ORIGIN

```

```

Query Match 26.2%; Score 131; DB 2; Length 157159;
Best Local Similarity 63.3%; Pred. No. 4.2e-23;
Matches 217; Conservative 0; Mismatches 125; Indels 1; Gaps 1;

QY 13 GAATTCATATCTCAAGTAGGAGATCTGTTATTCGCTGTTATTACATTTACACATAAG 72
|||||
Db 147633 GAAAAAATATTGAGACTCCCAAGACGACGAGATATTATACAGAAATATAAGAA 147692

QY 73 AAAGCTGAGGCTCTGAGAGGTCAAGATCAGCAGGTAAACAATAGCCCAAGACTTTGCT 132
|||||
Db 147693 AGTCCTCAAAATCAATCAGAAAAGAGAATAACAAGGAAAATGGCAAGGATATGAATA 147752

QY 133 TTAGAGCTGTGCTCTATCTGCTTTCTTTCTCCAAAAAACAACACTAATTTTGTGTGT 192
|||||
Db 147753 TGTGACTCACCCCAAAAACCCACTGGTTTGTCTCAACCTCACTAGTTTGTGTGTGT 147812

QY 193 TTTGTTTTTTTGTAGACAGGGTCTCGAGGTGTCAACCGAGGTGAGTGCAGTGCAGCGCA 252
|||||
Db 147813 TTTGTTTTTTTGTAGACAGAGTCTCACTGTTGCCACGGCTGAGTGCAGTGCAGCA 147872

QY 253 TTTGCACTCACCGCAACCTTCCGCTCC-CGCGTTTAAGCGATTTCTCTGCTCAGCTCCC 311
|||||
Db 147873 TCTCGGCTCACTGCACACCTCCGCTCTCTGGGTTAAAGTGAATTCCTGCTCAGCTCCC 147932

```

```

QY 312 AAGTAGCTGGGACTACAAAGCTCGGGACACACACGTAATAAATGAT 354
|||||
Db 147933 GAGTAGCTGGGACTACAGATGCCCGCCACACGCCCACTAAT 147975

```

RESULT 3

AC087525/c

```

LOCUS AC087525 157590 bp DNA linear PRI 14-MAY-2002
DEFINITION Homo sapiens chromosome 15, clone RP11-321G12, complete sequence.
ACCESSION AC087525
VERSION AC087525.6 GI:20564446
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 157590)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 15, clone RP11-321G12

```

REFERENCE

AUTHORS

TITLE

JOURNAL

AUTHORS

```

2 (bases 1 to 157590)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B., Brown,A.,
Camataata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
Collymore,A., Cooke,P., DeAvellino,K., Dewar,K., Diaz,J.S.,
Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J.,
Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T.,
Lehoczy,J., Levine,R., Liu,G., MacLean,C., Macdonald,P.,
Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rossetti,M.,
Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.

```

TITLE

JOURNAL

REFERENCE

AUTHORS

```

Submitted (06-JAN-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 157590)

```

```

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhgalter,B., Brown,A., Camataata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., DeAvellino,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gird,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rossetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

```

TITLE

JOURNAL

REFERENCE

AUTHORS

```

Submitted (02-MAY-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 157590)

```

```

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,

```


RESULT 5
AC017038
LOCUS
DEFINITION
20 unordered pieces.
AC017038
VERSION
HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 180707)
Waterston.R.H.
The sequence of Homo sapiens clone
Unpublished
REFERENCE
2 (bases 1 to 180707)
Waterston.R.H.
Direct Submission
Submitted (09-DEC-1999) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA

COMMENT

On Jun 16, 2000 this sequence version replaced gi:7230978.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH0313B02
----- Summary Statistics -----
Sequencing vector: M13; %
Sequencing vector: plasmid; %
Chemistry: Dye-primer ET; % of reads
Chemistry: Dye-terminator Big Dye; % of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 168786 bases at least Q40
Consensus quality: 172458 bases at least Q30
Consensus quality: 174496 bases at least Q20
Insert size: 200000; agarose-fp
Insert size: 178807; sum-of-contigs
Quality coverage: 3.48 in Q20 bases; agarose-fp
Quality coverage: 3.86 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 20 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 2076: contig of 2076 bp in length
* 2077 2176: gap of unknown length
* 2177 4769: contig of 2593 bp in length
* 4770 4869: gap of unknown length
* 4870 7558: contig of 2689 bp in length
* 7559 7659: gap of unknown length
* 7659 10945: contig of 3287 bp in length
* 10946 11045: gap of unknown length
* 11046 13945: contig of 2900 bp in length
* 13946 14046: gap of unknown length
* 14046 17724: contig of 3679 bp in length
* 17725 17824: gap of unknown length
* 17825 22156: contig of 4332 bp in length
* 22157 2257 26042: gap of unknown length
* 26043 26143: gap of unknown length
* 26143 32662: contig of 6520 bp in length
* 32663 32762: gap of unknown length
* 32763 41257: contig of 8495 bp in length
* 41258 43358 47568: contig of 6211 bp in length

* 47569 47668: gap of unknown length
* 47669 54391: contig of 6723 bp in length
* 54392 62210: gap of unknown length
* 62211 70009: contig of 7719 bp in length
* 70010 70109: gap of unknown length
* 70110 78968: contig of 8859 bp in length
* 78969 90881: gap of unknown length
* 90882 105415: contig of 11813 bp in length
* 105416 105516: gap of unknown length
* 105517 121842: contig of 16327 bp in length
* 121843 121943: gap of unknown length
* 121944 141937: contig of 19994 bp in length
* 141938 142037: gap of unknown length
* 142038 180707: contig of 38671 bp in length.
Location/Qualifiers
1. 180707
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/clone="RP11-313B2"

FEATURES

source

misc_feature
1. 2076
/note="assembly_name:Contig8"
misc_feature
2177. 4769
/note="assembly_name:Contig9"
misc_feature
4870. 7558
/note="assembly_name:Contig10"
misc_feature
7659. 10945
/note="assembly_name:Contig11"
misc_feature
11046. 13945
/note="assembly_name:Contig12"
misc_feature
14046. 17724
/note="assembly_name:Contig13"
misc_feature
17825. 22156
/note="assembly_name:Contig14"
misc_feature
22257. 26042
/note="assembly_name:Contig15"
misc_feature
26143. 32662
/note="assembly_name:Contig16"
misc_feature
32763. 41257
/note="assembly_name:Contig17"
misc_feature
41358. 47568
/note="assembly_name:Contig18"
misc_feature
47669. 54391
/note="assembly_name:Contig19"
misc_feature
54492. 62210
/note="assembly_name:Contig20"
misc_feature
62311. 70009
/note="assembly_name:Contig21"
misc_feature
70110. 78968
/note="assembly_name:Contig22"
misc_feature
79069. 90881
/note="assembly_name:Contig23"
misc_feature
90982. 105415
/note="assembly_name:Contig24"
clone_end.SP6
vector_side:right"
105516. 121842
/note="assembly_name:Contig25"
misc_feature
121943. 141936
/note="assembly_name:Contig26"
misc_feature
142037. 180707
/note="assembly_name:Contig27"
BASE COUNT 51636 a 38060 c 37973 g 51122 t 1916 others
ORIGIN

Query Match 25.7%; Score 128.6; DB 2; Length 180707;

Best Local Similarity 65.5%; Pred. No. 1.7e-22; Indels 0; Gaps 0;
Matches 188; Conservative 0; Mismatches 99;

```

Qy 90 AGGTCAAGATCAGCAGCTAACAAATGAGCAAGACTCTTGCTTTAGAGCTTGCTCTCTA 149
Db 174248 ATGTGACTATCAGTTAAACAATAGAAAGTATTTCACATATTTGTAGTATGTCACAAA 174307
Qy 150 TTCTTGCTTTCTTCCAAAAACACTACAAATTTTGTGTTTGTGTTTGTGTTGAGA 209
Db 174308 AACTTACAAAAACATTTACCAAGCAACACATCATCTTTTTTTTTTTTTTTTTTTGAGA 174367
Qy 210 CAGGCTTCGAGGTGTCACCCAGGCTGGAGTCAGTGGCGGATTCGACTCACCGCAAC 269
Db 174368 CAGAGTCTCGCTCTGTACCCAGGCTGGAGTGCATGTCATGCTCACTGCAAG 174427
Qy 270 CTCGCCCTCCGGCTTAAGCGATTCTCTGCTCAGCTCCCAAGTAGTGGGACTACAA 329
Db 174428 TTCCGCCCTCCGGCTTCTGCCATTCTCTGCTCAGCTCCCAAGTAGTGGGACTACAG 174487
Qy 330 GTCGGGACACCACTGAAAAATGATCAAGTTCTTAACATGATGATGCATA 376
Db 174488 GCACCTGCCACCAACCCAGCTAAATTTTTTTGTATATTTTTTAGTA 174534

RESULT 6
AC060764
LOCUS      184338 bp      DNA      linear      PRI 21-AUG-2002
DEFINITION Homo sapiens chromosome 8, clone RP11-626A5, complete sequence.
ACCESSION  AC060764
VERSION    AC060764.13  GI:22382034
KEYWORDS   HTG.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 184338)
            Birren,B., Nusbaum,C. and Lander,E.
            Homo sapiens chromosome 8, clone RP11-626A5
            Unpublished
            2 (bases 1 to 184338)
            Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczkzy,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivaz,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodors,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
            Direct Submission
            Submitted (20-APR-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
            3 (bases 1 to 184338)
            Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collamore,A.,
            Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Fato,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
            Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
            Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
            Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Rouse,C., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
            Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
            Tesfaye,S., Theodors,J., Topham,K., Travers,M., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
            Zembek,L., Zimmer,A. and Zody,M.
            Direct Submission
            Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
            4 (bases 1 to 184338)
            Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collamore,A.,
            Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Fato,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
            Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
            Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
            Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Rouse,C., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
            Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
            Tesfaye,S., Theodors,J., Topham,K., Travers,M., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
            Zembek,L., Zimmer,A. and Zody,M.
            Direct Submission
            Submitted (21-AUG-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
            On Aug 21, 2002 this sequence version replaced gi:22123659.
            All repeats were identified using RepeatMasker:
            Smit, A.P.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            ----- Genome Center
            Center: Whitehead Institute/ MIT Center for Genome Research
            Center code: WIBR
            Web site: http://www.seq.wi.mit.edu
            Contact: sequence_submissions@genome.wi.mit.edu
            ----- Project Information
            Center project name: L6260
            Center clone name: 626_A_5
            -----
            Location/Qualifiers
            1. 184338
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="8"
            /map="8"
            /clone="RP11-626A5"
            /clone_lib="RPCI-11 Human Male BAC"
            38. 87
            /rpt_family="WIR"
            /complement(167 417)
            /rpt_family="L1MA9"
            /complement(566 760)
            /rpt_family="L1MC5"
            877. 932
            /rpt_family="WIR"
            /complement(937 1312)
            /rpt_family="MLT1B"
            1391. 1480
            /rpt_family="WIR"
            1481. 1794
            /rpt_family="AluSq"
            1919. 1926
            /note="<30 qual SNGL region"
            1947. 1951
            /note="<30 qual SNGL region"
            1978. 2004
            /note="single clone coverage"

```



```

repeat_region complement(2310..2617)
repeat_region /rpt_family="AluY"
repeat_region complement(2854..3158)
unsure /rpt_family="AluJb"
repeat_region complement(3899..3904)
repeat_region /note="<30 qual SINGL region"
repeat_region complement(3934..4089)
unsure /rpt_family="FRAM"
unsure complement(3971..3976)
unsure /note="<30 qual SINGL region"
unsure complement(4059..4063)
unsure /note="<30 qual SINGL region"
unsure complement(4122..4128)
unsure /note="<30 qual SINGL region"
unsure complement(4125)
repeat_region /note="Probably C, possibly A"
repeat_region 4640..4933
repeat_region /rpt_family="AluSg"
repeat_region complement(5011..5494)
repeat_region /rpt_family="MLT1L"
repeat_region complement(5802..5971)
repeat_region /rpt_family="MIR"
repeat_region 6791..6935
repeat_region /rpt_family="MER3"
repeat_region 6936..7241
repeat_region /rpt_family="AluSg"
repeat_region complement(8712..8915)
repeat_region /rpt_family="MIR"
repeat_region 8952..9118
repeat_region /rpt_family="MLT1B"
repeat_region 9119..9331
repeat_region /rpt_family="AluJo"
repeat_region 9360..9560
repeat_region /rpt_family="MLT1B"
repeat_region complement(9680..9801)
repeat_region /rpt_family="MER96B"
repeat_region 9802..10090
repeat_region /rpt_family="AluSc"
repeat_region complement(10091..10117)
repeat_region /rpt_family="MER96B"
repeat_region complement(10146..10355)
repeat_region /rpt_family="AluY"
repeat_region 10436..10730
repeat_region /rpt_family="AluSx"
repeat_region complement(10745..10865)
repeat_region /rpt_family="L2"
repeat_region 11163..11304
repeat_region /rpt_family="L2"
repeat_region 11305..11342
repeat_region /rpt_family="(TTTAG)n"
repeat_region complement(11343..11622)
repeat_region /rpt_family="AluJb"
repeat_region 11623..11768
repeat_region /rpt_family="L2"
repeat_region 11859..12212
repeat_region /rpt_family="THE1B"
repeat_region complement(12223..12526)
repeat_region /rpt_family="MER110"
repeat_region 12580..12707
repeat_region /rpt_family="L2"
repeat_region complement(12855..13021)
repeat_region /rpt_family="MIR"
repeat_region 13047..13466
repeat_region /rpt_family="LIMB3"
repeat_region 13467..13778
repeat_region /rpt_family="AluSg"
repeat_region 13780..13803
repeat_region /rpt_family="(TA)n"
repeat_region 13804..13854
repeat_region /rpt_family="LIMB3"
repeat_region 13881..14073
repeat_region /rpt_family="LIMB4"
repeat_region complement(14074..14373)

```

```

/rpt_family="AluJo"
repeat_region 14374..14663
repeat_region /rpt_family="LIMB4"
repeat_region 14669..15128

Query Match 25.7%; Score 128.6; DB 9; Length 184338;
Best Local Similarity 65.5%; Pred. No. 1.7e-22; Mismatches 99; Indels 0; Gaps 0;
Matches 188; Conservative 0;

Qy 90 AGGTCAAGATCAGCAGCTAAACAATAGCAAGACTCTTGCTTTAGAGCTTTGCTCTCTA 149
Db 2217 ATGTGACTATCAGTTAAACAATAAGAAAGTTATTTCACATTATTGTAGTAGTCCAAA 2276

Qy 150 TCTTGTCTTTCTTTCCAAAAACACTACAATTTTGTGTTTGTGTTTGTGTTTGTGGA 209
Db 2277 AACTTACAAAAACATTTTACCAAGCAACACATCAATCTTTTTTTTTTTTTTTTGTGA 2336

Qy 210 CAGGTCTCGAGGTCTACCCAGGCTGAGTGCAGTGGCGGATTTCCGACTCACGCCAAC 269
Db 2337 CAGAGTCTCGGTCTGTCCAGGCTGGAGTGCAATGATCTCAGCTCAGTCTCAGTCAAG 2396

Qy 270 CTCGCCCTCCGCGCTTAAAGCGATTCTCTGCTCAGCCTCCCAAGTAGTGGGACTACAA 329
Db 2397 TTCGCCCTCCGCGGTCTGCGCATCTCTGCTCAGCCTCCCAAGTAGTGGGACTACAG 2456

Qy 330 GCTCGGGACACACGTAATAATGATCAAGTCTCAACATGTATGCATA 376
Db 2457 GCACCTGCCAGCACACCCAGCTAATTTTTTTGTATATTTTTTAGTA 2503

RESULT 7
AC009771_2/c
WPCOMMENT
Sequence split into 4 fragments LOCUS AC009771 Accession AC009771
Fragment Name Begin End
AC009771_0 1 110000
AC009771_1 100001 210000
AC009771_2 200001 310000
AC009771_3 300001 353131
Continuation (3 of 4) of AC009771 from base 200001 (AC009771 Homo sapiens chromosome 12

Query Match 25.3%; Score 126.4; DB 2; Length 110000;
Best Local Similarity 77.0%; Pred. No. 6.4e-22; Mismatches 46; Indels 0; Gaps 0;
Matches 154; Conservative 0;

Qy 145 CTCATTCTTGCTTTCTTTCCAAAAACACTACAATTTTGTGTTTGTGTTTGTGTTTGT 204
Db 65251 CTCAGGATGATTTGGCACTTAACAAAAAGTTAGTTCTTTTTTTTTTTTTTTTTTTT 65192

Qy 205 TGAGACAGGCTCTCGAGGTGTCAACCCAGGCTGGAGTGGCGCGATTTTCGACTCACC 264
Db 65191 TTAGACAGAACCTTGCCCTGTCAACCCAGGCTGGAGTGGCAATGGCACGATATTGACTCACT 65132

Qy 265 GCAACCTCGGCTCGCGCTTAAGCGATTCTCTGCTCAGCTCCCAAGTAGTGGGAC 324
Db 65131 GCAACCTCGGCTCGAGGTTCGAAGATTTCTCTGCTCAGCTCCCGAGTAGTGGGAC 65072

Qy 325 TACAGCTCGGACACACACG 344
Db 65071 TACAGGCAAGCCACACG 65052

RESULT 8
AC012386_0
WPCOMMENT
Sequence split into 4 fragments LOCUS AC012386 Accession AC012386
Fragment Name Begin End
AC012386_0 1 110000
AC012386_1 100001 210000
AC012386_2 200001 310000
AC012386_3 300001 352304
LOCUS AC012386 352304 bp DNA linear HTG 24-AUG-2002
DEFINITION Homo sapiens chromosome 12 clone RP11-642P15, WORKING DRAFT

```

```
SEQUENCE, 24 unordered pieces.
AC012386
VERSION AC012386.36 GI:22474781
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 352304)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Albrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaxia,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brileva,M., Brown,M., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,
Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Joudah,S.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Kovar,C.,
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louissegh,H.,
Lorazo,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maneshwar,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenwo,S., Oguh,M., Okwuonu,G.,
Organyue,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I.,
Sodergren,E., Sonaite,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vazquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 352304)
Worley,K.C.
Direct Submission
Submitted (27-OCT-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 352304)
Worley,K.C.
Direct Submission
Submitted (24-AUG-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Aug 24, 2002 this sequence version replaced gi:22450376.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Drafting Center Code: BCM
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HMPZ
Center clone name: RP11-642P15
FEATURES
source
1. 352304
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12"
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembled program: Phrap; version 0.990329
Consensus quality: 342731 bases at least Q40
Consensus quality: 346690 bases at least Q30
Consensus quality: 349268 bases at least Q20
Estimated insert size: 370200; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
2303: contig of 2303 bp in length
2403: gap of unknown length
2404
5961: contig of 3558 bp in length
6061: gap of unknown length
9865: contig of 3804 bp in length
9865: gap of unknown length
12751: contig of 2786 bp in length
12851: gap of unknown length
12852
15853: gap of unknown length
15854
1977: contig of 4124 bp in length
1978
2007: gap of unknown length
25949: contig of 5872 bp in length
26049: gap of unknown length
25950
33145: contig of 7096 bp in length
33146
33245: gap of unknown length
33246
40389: contig of 7144 bp in length
40390
40489: gap of unknown length
49662: contig of 9173 bp in length
49663
59828: contig of 10066 bp in length
59829
59928: gap of unknown length
70187
70286: gap of unknown length
81523: contig of 11237 bp in length
81524
89619: contig of 7996 bp in length
89620
89719: gap of unknown length
89720
105265: contig of 15546 bp in length
105266
105366
116343: contig of 10978 bp in length
116344
127455: contig of 11012 bp in length
127456
127555: gap of unknown length
143190: contig of 15635 bp in length
143191
143290: gap of unknown length
179464: contig of 36174 bp in length
179465
179565
230543: contig of 50979 bp in length
230544
230644: gap of unknown length
253492: contig of 22849 bp in length
253493
281092: gap of unknown length
281093
316138: contig of 27500 bp in length
316139
316238: gap of unknown length
316239
352304: contig of 36066 bp in length.
Location/Qualifiers
1. 352304
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12"
```

```

/clone="RP11-642B15"
BASE COUNT   99491 a 73601 c 7327 g 101345 t   4540 others
ORIGIN

Query Match      25.3% ; Score 126.4 ; DB 2 ; Length 110000 ;
Best Local Similarity 77.0% ; Pred No. 6.4e-22 ;
Matches 154 ; Conservative 0 ; Mismatches 46 ; Indels 0 ; Gaps 0 ;

QY 145 CTCTATTCTTCTTTCTTTCTTCCAAAAACACTACAATTTTGTGTTTGTGTTTGTGTTT 204
DB 29642 CTCTAGGCATGATTGGCACTTAACAAAAGTTAGTCTCTTTTTTTTTTTTTTTTTTTT 29701

QY 205 TGACACAGGCTCTCGAGTGTCAACCCAGGCTGGAGTGCAGTGGCGCATTCGACTACCC 264
DB 29702 TTAGACAGAACCTTGCCTGTCAACCCAGGCTGGAGTGCATGGCAGCATATTGACTCACT 29761

QY 265 GCAACCTCCGCTCCGCGCTTAAGCGATTCTCTGCTCCAGCTCCAGCTCCCAAGTAGCTGGGAC 324
DB 29762 GCAACCTCCGCTTCAGGTTCAAGAGATTCTCTGCTTCAGCTCCCGAGTAGCTGGGAC 29821

QY 325 TACAAGCTCGGGACACCAACG 344
DB 29822 TACAGGCAAGCGCCACCAACG 29841

RESULT 9
CNS01RHN      148068 bp      DNA      linear      PRI 02-MAY-2001
LOCUS         Human chromosome 14 DNA sequence BAC R-66E7 of library RPCI-11 from
DEFINITION    chromosome 14 of Homo sapiens (Human), complete sequence.
ACCESSION     AL161871
VERSION       AL161871.6   GI:13990336
KEYWORDS      HTG.
SOURCE        human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 148068)
AUTHORS       Heilig,R., Pettit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,
              Brottier,P., Cattolico,L., Barbe,V., Pelletier,E., Artiguenave,F.,
              Levy,M., Eckenberg,R., Bruls,T., Deberardinis,V., Craud,C.,
              Gysay,G., Saurin,W. and Weissbach,J.
              Sequencing of the human chromosome 14
              Unpublished
              2 (bases 1 to 148068)
              Genoscope.
              Direct Submission
              Submitted (02-MAY-2001) Genoscope - Centre National de Sequences :
              BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
              - Web : www.genoscope.cns.fr)
              On May 7, 2001 this sequence version replaced gi:13509291.
COMMENT       ----- Genome Center
              Center: Genoscope / Centre National de Sequences
              Center code: GS
              Web site: http://www.genoscope.cns.fr/
              Contact: SeqRef@genoscope.cns.fr
              -----
              The following BAC sequence is oriented from the T7 to the SP6 end.
              Upstream BAC (overlapping the T7 end) : R-603H7 (AC=AL109847)
              Downstream BAC (overlapping the SP6 end) : R-382E15 -----
              Summary Statistics
              Assembly program: Phrap; version 2.0
              Quality coverage: 8.50x in Q20 bases; sum-of-contigs
              -----
              Overall quality chart :
              Range : bases
              0 :
              1 - 9 :
              10 - 19 : 52
              20 - 29 : 261
              30 - 39 : 1435
              40 - 49 : 6761

```


only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
 This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr1>
 RP11-396D18 is from the library RP11-11.2 constructed by the group of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
 VECTOR: pBACEj.6

IMPORTANT: This sequence is not the entire insert of clone RP11-396D18. It may be shorter because we sequence overlapping sections only once, except for a short overlap.
 The true left end of clone RP11-175B9 is at 72090 in this sequence.
 The true right end of clone RP11-572A16 is at 2000 in this sequence.

```

FEATURES             Location/Qualifiers
     source            1..74089
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP11-396D18"
                     /clone_lib="RP11-11.2"
BASE COUNT           20957 a 14221 c 14800 g 24111 t
ORIGIN
Query Match          25.1%; Score 125.6; DB 9; Length 74089;
Best Local Similarity 76.0%; Pred. No. 1e-21;
Matches 155; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 141 TGTCTCTATTCTGCTTTCTTCCAAAACACTACAAATTTTGTGTTGTTT 200
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8257 TGTATTGTTATCATCTACTGTATTAGAAACAACTGATTTTTTTTTTTT 8198

QY 201 GTTTTGAGACAGGCTCTGAGGTGTCACCCAGCTGGAGTGCAGTGGCGGATTCGACT 260
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8197 TTTTGGATGGAGTCTCGCTCTGTCTCCAGCTGGAGTGCAGTGGCTTCTGGCT 8138

QY 261 CACCCGACCTCCGCTCCGCGCTTAAGGATTCCTCTGCTCAGCTCCCAAGTAGCTG 320
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8137 CACTGCAACCTCCGCTCCCGCTCAAGCAATTCCTGCTCAGCTCCCGAGTAGCTG 8078

QY 321 GGACTACAGCTCGGGACACACACG 344
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8077 GGATTACAGCATGTGCCACCATG 8054

```

```

RESULT 14
LOCUS      CNS01DSC      208309 bp      DNA      linear      PRI 02-MAY-2001
DEFINITION Human chromosome 14 DNA sequence BAC C-2593121 of library CalTech-D
            from chromosome 14 of Homo sapiens (Human), complete sequence.
ACCESSION  AL121767
VERSION     AL121767.6 GI:13990321
KEYWORDS   HTG.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 208309)
AUTHORS     Heilig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,
            Brottier,P., Cattolico,L., Barbe,V., Pelletier,E., Artiguenave,F.,

```

Levy M., Eckenberg R., Bruls T., deBerardinis V., Craud C.,
 Gyapay G., Saurin W. and Weissenbach J.
 Sequencing of the human chromosome 14
 Unpublished
 REFERENCE
 2 (bases 1 to 208309)
 AUTHORS
 TITLE
 JOURNAL
 Direct Submission
 Submitted (02-MAY-2001) Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 On May 7, 2001 this sequence version replaced gi:12580631.

----- Genom Center
 Center: Genoscope / Centre National de Sequencage
 Center code: GS
 Web site: <http://www.genoscope.cns.fr/>
 Contact: SeqRef@genoscope.cns.fr

 The following BAC sequence is oriented from the T7 to the SP6 end.
 Upstream BAC (overlapping the T7 end) : R-1074O12 (AC=AL049828)
 Downstream BAC (overlapping the SP6 end) : R-34018 -----
 Summary Statistics
 Assembly program: Phrap; version 2.0
 Quality coverage: 5.83x in Q20 bases; sum-of-contigs

 Overall quality chart :
 Range : bases
 0 :
 1 - 9 : 14
 10 - 19 : 218
 20 - 29 : 587
 30 - 39 : 2543
 40 - 49 : 6299
 50 - 59 : 6490
 60 - 69 : 10270
 70 - 79 : 28018
 80 - 89 : 74691
 90 - 99 : 79179

```

-----
Percentage of bases with a quality value >= 40 : 98 %
FEATURES             Location/Qualifiers
     source            1..208309
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="14"
                     /clone="C-2593121"
                     /clone_lib="CalTech-D"
                     /notes="matching EMBL:G04409"
                     RHdb:RH53476
                     RHdb:RH3633
                     dbSTS:STS25812
                     Identified using the e-PCR software (G. Schuler)
BASE COUNT           70845 a 38028 c 36323 g 63113 t
ORIGIN
Query Match          25.1%; Score 125.6; DB 9; Length 208309;
Best Local Similarity 80.0%; Pred. No. 9.6e-22;
Matches 160; Conservative 0; Mismatches 39; Indels 1; Gaps 1;

```

```

QY 146 TCTATTCTGCTTTCTTCCAAAACACTACAAATTTTGTGTTGTTTGT 205
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 74302 TTTCTTATCTACATCTCTACCAAAATCACTACAAATTTTGTGTTTGT 74243

QY 206 GAGACAGGCTTCGAGGTGTCAACCCAGCTGGAGTGCAGTGGCGGATTCGACTCACC 265
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 74242 GAGACGGAGTCTCCCTGTGTCTCCCGAGCTGGAGTGCAGTGGCGGATTCGACTC 74183

QY 266 CAACCTCGCTCCCGCTT-AAGCGATTCTCTGCTCAGCTCCCAAGTAGCTGGAC 324
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 74182 CAAGCTCGCTCCCGAGTTCACGCCATTCTCTGCTCAGCTCCCAAGTAGCTGGAC 74123

QY 325 TACAAGCTCGGGACACACACG 344

```

```

Db 74122 TACAGGCACCTGCACCATG 74103
|||||
RESULT 15
AC073228
LOCUS
DEFINITION Homo sapiens chromosome 11 clone RP11-660B16, WORKING DRAFT
ACCESSION AC073228
VERSION 2
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 153185)
Waterston,R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 153185)
Waterston,R.H.
Direct Submission
Submitted (10-JUN-2000) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Jul 27, 2000 this sequence version replaced gi:8440040.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H NH0660B16
----- Summary Statistics -----
Sequencing vector: M13; 100%
Chemistry: Dye-primer ET; 100% of reads
Assembly: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 133295 bases at least Q40
Consensus quality: 138801 bases at least Q30
Consensus quality: 141663 bases at least Q20
Insert size: 191000; agarose-fp
Quality coverage: 2.76 in Q20 bases; sum-of-contigs
Quality coverage: 3.17 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 40 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
1
1118: contig of 1118 bp in length
1218: gap of unknown length
1219: contig of 1692 bp in length
2911: gap of unknown length
3011: contig of 1388 bp in length
4399: gap of unknown length
4499: contig of 1074 bp in length
5573: gap of unknown length
5673: contig of 1390 bp in length
7063: gap of unknown length
7163: contig of 1702 bp in length
8865: gap of unknown length
8965: contig of 1512 bp in length
10477: gap of unknown length
12553: contig of 1977 bp in length
12554: gap of unknown length
12654: contig of 3399 bp in length

FEATURES
Source
1. 153185
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/clone="RP11-660B16"
1. 1118
/note="assembly_name:Contig19"
1219. 2910
/note="assembly_name:Contig20"
3011. 4398
misc_feature
misc_feature
misc_feature
16053 16152: gap of unknown length
16153 18644: contig of 2492 bp in length
18645 18744: gap of unknown length
20549: contig of 1805 bp in length
20550 20649: gap of unknown length
20650 22489: contig of 1840 bp in length
22490 22589: gap of unknown length
22590 25620: contig of 3031 bp in length
25621 25721: gap of unknown length
25722 28331: contig of 2611 bp in length
28332 28431: gap of unknown length
28432 31684: contig of 3253 bp in length
31685 34445: contig of 3661 bp in length
34446 35545: gap of unknown length
35546 38112: contig of 2567 bp in length
38113 38212: gap of unknown length
38213 41759: contig of 3547 bp in length
41760 44859: gap of unknown length
44860 44973: contig of 3114 bp in length
44974 45073: gap of unknown length
45074 49062: contig of 3988 bp in length
49063 52768: contig of 3607 bp in length
52769 52868: gap of unknown length
52869 55849: contig of 2981 bp in length
55850 59672: contig of 3723 bp in length
59673 59773: contig of 2785 bp in length
59774 62558: gap of unknown length
62559 66220: contig of 3562 bp in length
66221 70085: contig of 3766 bp in length
70086 70186: gap of unknown length
70187 73879: contig of 3694 bp in length
73880 73979: gap of unknown length
73980 79836: contig of 5757 bp in length
79837 84227: contig of 4291 bp in length
84228 85206: contig of 4979 bp in length
85207 89306: gap of unknown length
89307 94087: contig of 4781 bp in length
94088 94188: gap of unknown length
94189 99446: contig of 5259 bp in length
99447 104310: contig of 4764 bp in length
104311 104410: gap of unknown length
104411 110222: contig of 5812 bp in length
110223 110322: gap of unknown length
110323 116064: contig of 5742 bp in length
116065 121387: contig of 5223 bp in length
121388 121488: gap of unknown length
121489 128499: gap of unknown length
128500 135537: contig of 6939 bp in length
135538 136637: gap of unknown length
136638 142005: contig of 6368 bp in length
142006 153185: contig of 11080 bp in length.
153186: Location/Qualifiers
1. 153185
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/clone="RP11-660B16"
1. 1118
/note="assembly_name:Contig19"
1219. 2910
/note="assembly_name:Contig20"
3011. 4398

```


GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 18:18:38 ; Search time 163 Seconds
(without alignments)
5565.337 Million cell updates/sec

Title: US-09-434-382-3
Perfect score: 2958
Sequence: 1 cgcggcgtaggtgaccggc.....ataaagattgagtttgcga 2958

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq:
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2958	100.0	2958	4	US-09-564-805-3
2	2874.4	97.2	2908	4	US-09-564-805-223
3	2819.6	95.3	2892	4	US-09-564-805-225
4	2481	83.9	2481	4	US-09-564-805-1
5	1645.6	55.6	2470	4	US-09-564-805-221
6	657.2	22.2	26664	4	US-09-564-805-28
7	655	22.1	655	4	US-09-564-805-27
8	297.4	10.1	350	4	US-09-564-805-210
9	295	10.0	295	4	US-09-564-805-4
10	237	8.0	238	4	US-09-328-111-315
11	145	4.9	145	4	US-09-564-805-26
12	139	4.7	139	4	US-09-564-805-16
13	139	4.7	139	4	US-09-564-805-20
14	121	4.1	121	4	US-09-564-805-24
15	120	4.1	120	4	US-09-564-805-10
16	119	4.0	119	4	US-09-564-805-18
17	118.4	4.0	326	4	US-09-564-805-212
18	113	3.8	113	4	US-09-564-805-14
19	110	3.7	110	4	US-09-564-805-22
20	100	3.4	100	4	US-09-564-805-23
21	97	3.3	97	4	US-09-564-805-19
22	96	3.2	96	4	US-09-564-805-15
23	86	2.9	86	4	US-09-564-805-17
24	79	2.7	79	4	US-09-564-805-25
25	73	2.5	73	4	US-09-564-805-13
26	71	2.4	71	4	US-09-564-805-6
27	69	2.3	69	4	US-09-564-805-9

28 65 2.2 65 4 US-09-564-805-7 Sequence 7, Appli
29 59 2.0 59 4 US-09-564-805-11 Sequence 11, Appli
30 59 2.0 59 4 US-09-564-805-12 Sequence 12, Appli
31 58.4 2.0 2517 4 US-09-315-794-51 Sequence 51, Appli
32 58.4 2.0 2517 4 US-09-389-341-51 Sequence 51, Appli
33 58 2.0 58 4 US-09-564-805-8 Sequence 8, Appli
34 51 1.7 51 4 US-09-564-805-5 Sequence 5, Appli
35 48.6 1.6 4411529 4 US-09-103-840A-1 Sequence 1, Appli
36 48.2 1.6 1926 4 US-09-249-585A-2 Sequence 2, Appli
37 48.2 1.6 2580 3 US-09-050-863-2 Sequence 2, Appli
38 48.2 1.6 2580 4 US-09-359-081-2 Sequence 2, Appli
39 48.2 1.6 5452 2 US-09-130-114-1 Sequence 1, Appli
40 48.2 1.6 9600 4 US-08-910-647-1 Sequence 1, Appli
41 48.2 1.6 9600 4 US-09-620-925-1 Sequence 1, Appli
42 48.2 1.6 10596 1 US-07-884-811-15 Sequence 15, Appli
43 48.2 1.6 10596 1 US-07-885-971-15 Sequence 15, Appli
44 48.2 1.6 10596 1 US-08-087-783A-15 Sequence 15, Appli
45 48.2 1.6 10596 1 US-08-194-088B-15 Sequence 15, Appli

ALIGNMENTS

RESULT 1
US-09-564-805-3
; Sequence 3, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigan, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: Gene and a Paralog and Orthologous Genes
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2958
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51)..(2531)
; OTHER INFORMATION: coding sequence as in SEQ ID NO:1
US-09-564-805-3

Query Match 100.0%; Score 2958; DB 4; Length 2958;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2958; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CGCGGGCTAGGTGACGGCGGCTTTCTCAGTTTGGTGGAGACGGCGGCATGTGGGCG 60
Db 1 CGCGGGCTAGGTGACGGCGGCTTTCTCAGTTTGGTGGAGACGGCGGCATGTGGGCG 60
Qy 61 TTTGTCGCTGCTGCGGTCCTCGCGGCGGACGCCATGTGCGAGGAGCGACCATATCGC 120
Db 61 TTTGTCGCTGCTGCGGTCCTCGCGGCGGACGCCATGTGCGAGGAGCGACCATATCGC 120
Qy 121 AGGACACCGCGCGCGCGAGCGCGCGCAAGAGACCCGCTGCGGCACTGCGCAGCGGAG 180
Db 121 AGGACACCGCGCGCGCGAGCGCGCGCAAGAGACCCGCTGCGGCACTGCGCAGCGGAG 180
Qy 181 AGAAGCGGGACCGTCGCGGGTGCTCCGGCGGCGCCCAACACCGTGACCTGCAGGTGGTGG 240
Db 181 AGAAGCGGGACCGTCGCGGGTGCTCCGGCGGCGCCCAACACCGTGACCTGCAGGTGGTGG 240

Qy 241 CAGCGGTAGCGGGACTCGGGCGCGGCTCTAGTCTTCTCCGAGTTCAACCGGTATC 300
Db 241 CAGCGGTAGCGGGACTCGGGCGCGGCTCTAGTCTTCTCCGAGTTCAACCGGTATC 300
Qy 301 TCTTCAACTGTGGAGAGCGGTTGACAGACTCATGACGAGACACAAGTTAAAGGTTGCTC 360
Db 301 TCTTCAACTGTGGAGAGCGGTTGACAGACTCATGACGAGACACAAGTTAAAGGTTGCTC 360
Qy 361 GCCTGGACAAATATCTCTGACAGAAATGCACTGCTCTAAATGTTGGGGGCTTAAAGTGAA 420
Db 361 GCCTGGACAAATATCTCTGACAGAAATGCACTGCTCTAAATGTTGGGGGCTTAAAGTGAA 420
Qy 421 TGATCTTACTTTAAAGAAACCGGGCTTCAAAGTGTGTAATTTCTGGAGCTCCACAAAC 480
Db 421 TGATCTTACTTTAAAGAAACCGGGCTTCAAAGTGTGTAATTTCTGGAGCTCCACAAAC 480
Qy 481 TGGAAAAATACCTCGAAGCAATCAAAATATTTCTGGTCCATTTGAAAGGAATAGAATGG 540
Db 481 TGGAAAAATACCTCGAAGCAATCAAAATATTTCTGGTCCATTTGAAAGGAATAGAATGG 540
Qy 541 CTGTGCGGCCCACTCTGCCCCAGAATACGAGGATGAAACCATGACAGTTTACCAGATCC 600
Db 541 CTGTGCGGCCCACTCTGCCCCAGAATACGAGGATGAAACCATGACAGTTTACCAGATCC 600
Qy 601 CCATACACAGTGAACAGAGGAGGGAAGCAACCACTGACAGATCCAGAAAGGCCTC 660
Db 601 CCATACACAGTGAACAGAGGAGGGAAGCAACCACTGACAGATCCAGAAAGGCCTC 660
Qy 661 TCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCCGAATGAAATGAGCCACACC 720
Db 661 TCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCCGAATGAAATGAGCCACACC 720
Qy 721 TTCCACATGGTGTAGCCAGAGAAGAGGGGTCAAGGACTCTTCCCTGGTCTGATCTTTCA 780
Db 721 TTCCACATGGTGTAGCCAGAGAAGAGGGGTCAAGGACTCTTCCCTGGTCTGATCTTTCA 780
Qy 781 TCTGTAAGCTTCACTTAAAGAGAGAACTTCTGGTCTCAAAGCAAGAGAGATGGGCC 840
Db 781 TCTGTAAGCTTCACTTAAAGAGAGAACTTCTGGTCTCAAAGCAAGAGAGATGGGCC 840
Qy 841 TCCAGTTGGGACAGCTGCCATCGCTCCCATCTTCTGCTCAAGGACGGGAAAGCA 900
Db 841 TCCAGTTGGGACAGCTGCCATCGCTCCCATCTTCTGCTCAAGGACGGGAAAGCA 900
Qy 901 TCATCTATGAAGAGAGAGATTTGGTGTGAAGACTGTGACTCTCTCCAGATCCTGGTG 960
Db 901 TCATCTATGAAGAGAGAGATTTGGTGTGAAGACTGTGACTCTCTCCAGATCCTGGTG 960
Qy 961 CTGCTTTTGGTGTAGAAATGTCCAGATGAAGCTTCAITTCACCCCATCTGTGAGAAATG 1020
Db 961 CTGCTTTTGGTGTAGAAATGTCCAGATGAAGCTTCAITTCACCCCATCTGTGAGAAATG 1020
Qy 1021 CCACCTTTTCAAGGATACCAAGAAAGGAGATGCCCGCTGGCTTGGTGTTCACATGG 1080
Db 1021 CCACCTTTTCAAGGATACCAAGAAAGGAGATGCCCGCTGGCTTGGTGTTCACATGG 1080
Qy 1081 CCCAGCATCTGTCTGTGACAGCAGGTATCCAGCAGTGGATGGAGAGTTTGGGCTG 1140
Db 1081 CCCAGCATCTGTCTGTGACAGCAGGTATCCAGCAGTGGATGGAGAGTTTGGGCTG 1140
Qy 1141 ACACCCAGCACTTGGTCTGTAATGAACTGTGCTCAGTTTCAACCTTCGACGCCACA 1200
Db 1141 ACACCCAGCACTTGGTCTGTAATGAACTGTGCTCAGTTTCAACCTTCGACGCCACA 1200
Qy 1201 AGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCTGTCTACCAAGTTTCC 1260
Db 1201 AGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCTGTCTACCAAGTTTCC 1260
Qy 1261 GCTGTAAAGAGAGGGCCCCACCTCAGTGTGCCCATGGTTACAGGTTGAATGCCTCTCA 1320
Db 1261 GCTGTAAAGAGAGGGCCCCACCTCAGTGTGCCCATGGTTACAGGTTGAATGCCTCTCA 1320

Qy 1321 AGTACCAGCTCCGTTCCAGGAGGAGTGGCAGAGGATGCATTATTACTTTCGAATCCTG 1380
Db 1321 AGTACCAGCTCCGTTCCAGGAGGAGTGGCAGAGGATGCATTATTACTTTCGAATCCTG 1380
Qy 1381 AGGAATTCATAGTTGAGCGCTGTCAGCTTCCCAACTTTCACGACGAGCGTGCAGGATACA 1440
Db 1381 AGGAATTCATAGTTGAGCGCTGTCAGCTTCCCAACTTTCACGACGAGCGTGCAGGATACA 1440
Qy 1441 GGAGGAGTGGCAGGACGGCCCGACCCAGCAGAGAAAAGAGTCACTACCCAGAAATCA 1500
Db 1441 GGAGGAGTGGCAGGACGGCCCGACCCAGCAGAGAAAAGAGTCACTACCCAGAAATCA 1500
Qy 1501 TCTTCTTTGAAACAGGCTCTGCCATCCCGATGAAGATTCGAAATGTTCAGTGCACACTTG 1560
Db 1501 TCTTCTTTGAAACAGGCTCTGCCATCCCGATGAAGATTCGAAATGTTCAGTGCACACTTG 1560
Qy 1561 TCAACATAAGCCCGCACACGCTCTCTGCTACTGGACTGTGGTGGGGCACATTTGGGGCAGC 1620
Db 1561 TCAACATAAGCCCGCACACGCTCTCTGCTACTGGACTGTGGTGGGGCACATTTGGGGCAGC 1620
Qy 1621 TGTGCCGTCTATTACGGAGACAGGTGGACAGGGTCTCTGGGACCCCTGGCTGCTGTGTTG 1680
Db 1621 TGTGCCGTCTATTACGGAGACAGGTGGACAGGGTCTCTGGGACCCCTGGCTGCTGTGTTG 1680
Qy 1681 TGTCCCACTGACGAGATCACACACGGCTTGCCCAAGTATCTTCTGCAGAGAGAAC 1740
Db 1681 TGTCCCACTGACGAGATCACACACGGCTTGCCCAAGTATCTTCTGCAGAGAGAAC 1740
Qy 1741 GCGCTTTGGCATCTTTGGGAAAGCCGCTTCAACCTTTTCTGGTGGTTCGCCCAACCCAGC 1800
Db 1741 GCGCTTTGGCATCTTTGGGAAAGCCGCTTCAACCTTTTCTGGTGGTTCGCCCAACCCAGC 1800
Qy 1801 TCAAGCCTGGCTCCAGCAGTACCAACACAGTGCAGAGGGTCTTCACACACATCAGTA 1860
Db 1801 TCAAGCCTGGCTCCAGCAGTACCAACACAGTGCAGAGGGTCTTCACACACATCAGTA 1860
Qy 1861 TGATTCCTGCCAAATGCTTTCAGNAAGGGCTGAGATCTCCAGTCTCCAGTGGAAAGAT 1920
Db 1861 TGATTCCTGCCAAATGCTTTCAGNAAGGGCTGAGATCTCCAGTCTCCAGTGGAAAGAT 1920
Qy 1921 TGATCAGTTCCGCTTTCGGAACATGTATTTGGAAGAGTTTCAGACCTGTCTGGTGGCGC 1980
Db 1921 TGATCAGTTCCGCTTTCGGAACATGTATTTGGAAGAGTTTCAGACCTGTCTGGTGGCGC 1980
Qy 1981 ACTGCAAGCATGGCTTGGCTGTGCGCTGTGCACACCTCTGGCTGGAAGTGTCTATT 2040
Db 1981 ACTGCAAGCATGGCTTGGCTGTGCGCTGTGCACACCTCTGGCTGGAAGTGTCTATT 2040
Qy 2041 CCGGGACACCATGCGCTCGAGGCTCTGTCGGATGGGAAAGATGCCACCTCTCTGA 2100
Db 2041 CCGGGACACCATGCGCTCGAGGCTCTGTCGGATGGGAAAGATGCCACCTCTCTGA 2100
Qy 2101 TACATGAAGCCACCTTGGAAAGATGGTTTGAAGAGAGCAGTGGAAAGACACACAGCA 2160
Db 2101 TACATGAAGCCACCTTGGAAAGATGGTTTGAAGAGAGCAGTGGAAAGACACACAGCA 2160
Qy 2161 CAACGTCCTCAAGCCATCAGCTGGGATCGGATGAACGCGGAGTTTCATTATGCTGAACC 2220
Db 2161 CAACGTCCTCAAGCCATCAGCTGGGATCGGATGAACGCGGAGTTTCATTATGCTGAACC 2220
Qy 2221 ACTTCAGCAGCGTATGCCAAGTCCCTCTTTCAGCCCCCACTTCAGCGAGAAAGTGG 2280
Db 2221 ACTTCAGCAGCGTATGCCAAGTCCCTCTTTCAGCCCCCACTTCAGCGAGAAAGTGG 2280
Qy 2281 GAGTTGCTTTTGAACCATGAAGTCTGCTTTTGGAGACTTTCACCAATGCCCAAGCTGA 2340
Db 2281 GAGTTGCTTTTGAACCATGAAGTCTGCTTTTGGAGACTTTCACCAATGCCCAAGCTGA 2340
Qy 2341 TTCCCCCAGCTGAAGCCCTGTTTGTGGCGACATCGAGGAGATGGAGGCGGAGGAGA 2400
Db 2341 TTCCCCCAGCTGAAGCCCTGTTTGTGGCGACATCGAGGAGATGGAGGCGGAGGAGA 2400
Qy 2401 AGCGGAGCTGCGCAGGTGCGGGCGGCCCTCTCTGTCCAGGAGCTGGCAGGCGCCTGG 2460

```
Db 2401 AGCGGAGCTGCGGAGGTCGCGGCGGCCCTCTCTGTCCAGGAGCTGGCAGGCGCCTGG 2460
Qy 2461 AGGATGGGAGGCTCAGCAGAAAGCGGGCCACACAGAGAGGACACAGGCAAGAAAGGTCA 2520
Db 2461 AGGATGGGAGGCTCAGCAGAAAGCGGGCCACACAGAGAGGACACAGGCAAGAAAGGTCA 2520
Qy 2521 GAGCCAGTGAAGATCTGGGAGACCTGAACTCAGAAAGCTGTGTCTTCTGCCCCACG 2580
Db 2521 GAGCCAGTGAAGATCTGGGAGACCTGAACTCAGAAAGCTGTGTCTTCTGCCCCACG 2580
Qy 2581 CACGACACCGTATCTGCCCTCTCTGTGTGTAGAGCTGAAGAGCAGGTCGCCCAGGAGG 2640
Db 2581 CACGACACCGTATCTGCCCTCTCTGTGTGTAGAGCTGAAGAGCAGGTCGCCCAGGAGG 2640
Qy 2641 CAGCTCAGGATAGGTGTGTAGGAGCTGTGCCGAGGCTTGGGCTCCACATAAGCACTAGT 2700
Db 2641 CAGCTCAGGATAGGTGTGTAGGAGCTGTGCCGAGGCTTGGGCTCCACATAAGCACTAGT 2700
Qy 2701 CTATAGATGCTCTTAGGACTGGTGTGGACAGCGCGGGCCAGGAGGCTGCCACAGG 2760
Db 2701 CTATAGATGCTCTTAGGACTGGTGTGGACAGCGCGGGCCAGGAGGCTGCCACAGG 2760
Qy 2761 GAAGCAAGCAGATGAACATAATTTTCAAGGCAAGTCTTTTAAAGAGTCTTTGGAAACAG 2820
Db 2761 GAAGCAAGCAGATGAACATAATTTTCAAGGCAAGTCTTTTAAAGAGTCTTTGGAAACAG 2820
Qy 2821 ACGGGGCAACCTTTCTCTTAATCCAGCAAGTGATTCCTCGCACACACAGAGCAAGCAGA 2880
Db 2821 ACGGGGCAACCTTTCTCTTAATCCAGCAAGTGATTCCTCGCACACACAGAGCAAGCAGA 2880
Qy 2881 GTAACAGGATCAGTGGGTCTAAGTGTCCAGACTTAACGAAATAGTATTCAGCTGCAA 2940
Db 2881 GTAACAGGATCAGTGGGTCTAAGTGTCCAGACTTAACGAAATAGTATTCAGCTGCAA 2940
Qy 2941 TAAAGATTGAGTTTGCAA 2958
Db 2941 TAAAGATTGAGTTTGCAA 2958
```

RESULT 2

```
US-09-564-805-223
; Sequence 223, Application US/09564805
; Patent No. 633403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; PRIORITY FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 223
; LENGTH: 2908
; TYPE: DNA
; ORGANISM: Pan troglodytes
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
US-09-564-805-223
```

Query Match 97.2%; Score 2874.4; DB 4; Length 2908;

Best Local Similarity 99.3%; Pred. No. 0; Mismatches 21; Indels 0; Gaps 0;

```
Qy 51 ATGTGGGCGCTTTGCTCGCTCTGCGTCCGCGGCCGACGACCATTCGAGGAGCGC 110
Db 1 ATGTGGGCGCTTTGCTCGCTCTGCGTCCGCGGCCGACGACCATTCGAGGAGCGC 60
Qy 111 ACCATATCGCAGGACACCGCGCGCGAGCGCGGCAAGGACCGCTCGGCGACCTG 170
Db 61 ACCATATCGCAGGACACCGCGCGCGAGCGCGGCAAGGACCGCTCGGCGACCTG 120
Qy 171 CGCAGCGAGAGAGCGCGGACCGTCCGGGTGCTCCGCGGCCCAAAACACCGTGTACCTG 230
Db 121 CGCAGCGAGAGAGCGCGGACCGTCCGGGTGCTCCGCGGCCCAAAACACCGTGTACCTG 180
Qy 231 CAGGTGTGGCAGCGGCTAGCCGGGACTCGGGCGCGCGCTCTAGCTTCTTCCAGTTC 290
Db 181 CAGGTGTGGCAGCGGCTAGCCGGGACTCGGGCGCGCGCTCTAGCTTCTTCCAGTTC 240
Qy 291 AACCGGTATCTCTCAACTGTGGAGAGGCAATCAGAGACTCATGCGAGGACACAAGTTA 350
Db 241 AACCGGTATCTCTCAACTGTGGAGAGGCAATCAGAGACTCATGCGAGGACACAAGTTA 300
Qy 351 AAGGTGTGCTCGCCTGGGCAACATATTCCTGACAGCAATGCACTGTCTAATGTTGGGGC 410
Db 301 AAGGTGTGCTCGCCTGGGCAACATATTCCTGACAGCAATGCACTGTCTAATGTTGGGGC 360
Qy 411 TTAAGTGAATGATTTCTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGA 470
Db 361 TTAAGTGAATGATTTCTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGA 420
Qy 471 CCTCCAACTCGGAAATACCTCGAGCAATCAAAATATTTCTGGTCCATTTGAAGA 530
Db 421 CCTCCAACTCGGAAATACCTCGAGCAATCAAAATATTTCTGGTCCATTTGAAGA 480
Qy 531 ATAGAACTGGGTGTGGCGGCCCACTCTGCCAGAAATACGAGGATGAACCATGACAGTT 590
Db 481 ATAGAACTGGGTGTGGCGGCCCACTCTGCCAGAAATACGAGGATGAACCATGACAGTT 540
Qy 591 TACCAGATCCCCATACACAGTGAACAGAGGAGGAAAGCAACCATGCGAGTCCA 650
Db 541 TACCAGATCCCCATACACAGTGAACAGAGGAGGAAAGCAACCATGCGAGTCCA 600
Qy 651 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGGATCTTCAGACTCCGAGTCGAATGA 710
Db 601 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGGATCTTCAGACTCCGAGTCGAATGA 660
Qy 711 GAGCCACACCTTCCACATGGTGTAGCCAGAGAGGAGGCTCAGGAGCTCTTCCCTGGTC 770
Db 661 GAGCCACACCTTCCACATGGTGTAGCCAGAGAGGAGGCTCAGGAGCTCTTCCCTGGTC 720
Qy 771 GTAGCTTTTCTGTAAAGCTTCACTTAAAGAGAGGAAACTTCTTGGTGTCTCAAGCAAG 830
Db 721 GTAGCTTTTCTGTAAAGCTTCACTTAAAGAGAGGAAACTTCTTGGTGTCTCAAGCAAG 780
Qy 831 GAGATGGGCTCCAGTGGGACAGCTCCATCGCTCCATCATTCGCTGTCTCAAGAC 890
Db 781 GAGATGGGCTCCAGTGGGACAGCTCCATCGCTCCATCATTCGCTGTCTCAAGAC 840
Qy 891 GGGAAAGCATCACTCATGAAGAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTCCA 950
Db 841 GGGAAAGCATCACTCATGAAGAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTCCA 900
Qy 951 GATCCTGTGTCTTTTGTGGTGTAGAAATGTCCAGATGAAGGCTTCATTCAACCCATC 1010
Db 901 GATCCTGTGTCTTTTGTGGTGTAGAAATGTCCAGATGAAGGCTTCATTCAACCCATC 960
Qy 1011 TGTGAGATGCCCTTTTTCAGAGTACCAAGAAAGGAGATGCCCGCTGGCTTGGTG 1070
Db 961 TGTGAGATGCCCTTTTTCAGAGTACCAAGAAAGGAGATGCCCGCTGGCTTGGTG 1020
Qy 1071 GTTCACATGGGCGCCAGCATCTGTCTGTGACAGAGGATCCAGCAGTGGATGAGAGG 1130
Db 1021 GTTCACATGGGCGCCAGAAATCTGTCTTGTGACAGAGGATCCAGCAGTGGATGAGAGG 1080
```

Qy 1131 TTTGGGCTGACACCCAGCAGCTTGGTCTGATGAGACTGTGCTCAGTTCAACCTT 1190
 Db 1081 TTTGGGCTGACACCCAGCAGCTTGGTCTGATGAGACTGTGCTCAGTTCAACCTT 1140
 Qy 1191 CGCAGCCACAAGATTCAAAACCCAGCTCAACCTCATCCACCCGAGCATCTTCCCTCGTCTC 1250
 Db 1141 CGCAGCCACAAGATTCAAAACCCAGCTCAACCTCATCCACCCGAGCATCTTCCCTCGTCTC 1200
 Qy 1251 ACCAGTTTCCGCTGTAAGAGAGGAGGCCCCACCTCTAGTGTGCCATGTTGAGGGTAA 1310
 Db 1201 ACCAGTTTCCGCTGTAAGAGAGGAGGCCCCACCTCTAGTGTGCCATGTTGAGGGTAA 1260
 Qy 1311 TGCCTCTCAAGTACCACTCCGTCAGGAGGAGTGGCAGAGGATGCCATTATTA 1370
 Db 1261 TGCCTCTCAAGTACCACTCCGTCAGGAGGAGTGGCAGAGGATGCCATTATTA 1320
 Qy 1371 TGCATCTGAGGAATTCATAGTTGAGGCGCTGCACTTCCCACTCCAGCAGAGCGTG 1430
 Db 1321 TGCATCTGAGGAATTCATAGTTGAGGCGCTGCACTTCCCACTCCAGCAGAGTG 1380
 Qy 1431 CAGAGTACAGAGAGTGCAGAGGAGCGCCAGCCAGCCAGAGAGAAAAGTCACTAC 1490
 Db 1381 CAGAGTACAGAGAGTGCAGAGGAGCGCCAGCCAGCCAGAGAGAAAAGTCACTAC 1440
 Qy 1491 CCAGAAATCATCTTCTTGGACAGGCTTGCATCCCGATCCGATGAGATTCGAATGTCAGT 1550
 Db 1441 CCAGAAATCATCTTCTTGGACAGGCTTGCATCCCGATCCGATGAGATTCGAATGTCAGT 1500
 Qy 1551 GCCACATCTGTCATTAAGCCCGACACGCTCTCTGCTACTTGGACTGTGGTGAAGGACACA 1610
 Db 1501 GCCACATCTGTCATTAAGCCCGACACGCTCTCTGCTACTTGGACTGTGGTGAAGGACAC 1560
 Qy 1611 TTTGGGAGCTGTGCCGTATTAACGAGACAGAGTGGACAGGCTCTGGGACCCCTGGCT 1670
 Db 1561 TTTGGGAGCTGTGCCGTATTAACGAGACAGAGTGGACAGGCTCTGGGACCCCTGGCT 1620
 Qy 1671 GCTGTGTTTGTGTCACCTGACACGATCACACAGGCTTGCACAGTATCTTGGTG 1730
 Db 1621 GCTGTGTTTGTGTCACCTGACACGATCACACAGGCTTGCACAGTATCTTGGTG 1680
 Qy 1731 CAGAGAGAACCGCTTGGCATCTTTGGGAAAGCCGCTTACCCCTTCTGCTGGTGTGCC 1790
 Db 1681 CAGAGAGAACCGCTTGGCATCTTTGGGAAAGCCCTTACCCCTTCTGCTGGTGTGCC 1740
 Qy 1791 CCCAACAGCTCAAGCTGTGCTCCAGCAGTACCAACACAGTGCAGAGGCTCTGCAC 1850
 Db 1741 CCCAACAGCTCAAGCTGTGCTCCAGCAGTACCAACACAGTGCAGAGGCTCTGCAC 1800
 Qy 1851 CACATCAGTATGATTCCTGCCAATGCCCTCAGGAGGGGCTGAGATCTCCAGTCCCTGCA 1910
 Db 1801 CACATCAGTATGATTCCTGCCAATGCCCTCAGGAGGGGCTGAGATCTCCAGTCCCTGCA 1860
 Qy 1911 GTGGAAGATTGATCAGTTCGCTGTGGCAACATGTGATTTGGAAGAGTTTTCAGACCTGT 1970
 Db 1861 GTGGAAGATTGATCAGTTCGCTGTGGCAACATGTGATTTGGAAGAGTTTTCAGACCTGT 1920
 Qy 1971 CTGGTGGGCACTCAAGATGCTGCTGTGGCTGTGGCTGACACCTCTGGCTGGAAA 2030
 Db 1921 CTGGTGGGCACTCAAGATGCTGCTGTGGCTGTGGCTGACACCTCTGGCTGGAAA 1980
 Qy 2031 GTGGTCTATTCGGGGACACATGCCCTCGGAGGCTCTGGTCCGATGGGAAAGATGCC 2090
 Db 1981 GTGGTCTATTCGGGGACACATGCCCTCGGAGGCTCTGGTCCGATGGGAAAGATGCC 2040
 Qy 2091 ACCCTCTGATACATGAAGCCACCTCGAAGATGTTTGGAGAGAGACAGTGGAAAAG 2150
 Db 2041 ACCCTCTGATACATGAAGCCACCTCGAAGAGCTTTTGGAGAGAGACAGTGGAAAAG 2100
 Qy 2151 ACACACAGACAACTCCCAAGCCATCAGCTGGGGATGCGGATGAACCGGAGTTCA 2210
 Db 2101 ACACACAGACAACTCCCAAGCCATCAGCTGGGGATGCGGATGAACCGGAGTTCA 2160
 Qy 2211 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGTCTCCCTCTTCAGCCCCCACTTCAGC 2270

Db 2161 ATGCTGAACCACTTCAGCCAGCGCTATGCAAGTCCCCCTCTTCAGCCCCCACTTCAC 2220
 Qy 2271 GAGAAAGTGGAGTGTCTTTTGACCACATGAAGTGTCTTTTGAGACTTTTCAACAATG 2330
 Db 2221 GAGAAAGTGGAGTGTCTTTTGACCACATGAAGTGTCTTTTGAGACTTTTCAACAATG 2280
 Qy 2331 CCCAAGCTGATTTCCCACTGAAAGCCCTGTTTGTCTGGCCACATCGAGGAGATGGAGG 2390
 Db 2281 CCCAAGCTGATTTCCCACTGAAAGCCCTGTTTGTCTGGCCACATCGAGGAGATGGAGG 2340
 Qy 2391 CGCAGGAGAAAGCGGAGCTGCGGAGTCCGGCGGCTCTCTGTCCAGGAGCTGGCA 2450
 Db 2341 CGCAGGAGAAAGCGGAGCTGCGGAGTCCGGCGGCTCTCTGTCCAGGAGCTGGCA 2400
 Qy 2451 GCGGCTCTGGAGTGGGAGCTTCAGCAAGAACCGGCCCCACACAGAGGAGCCACAGGCC 2510
 Db 2401 GCGGCTCTGGAGTGGGAGCTTCAGCAAGAACCGGCCCCACACAGAGGAGCCACAGGCC 2460
 Qy 2511 AAGAGTCAAGGCCAGTGAAGATCTGGGAGACCTTGAATCTCAGAAGGCTGTGTCTT 2570
 Db 2461 AAGAGTCAAGGCCAGTGAAGATCTGGGAGACCTTGAATCTCAGAAGGCTGTGTCTT 2520
 Qy 2571 CTGCCCCACACGACCGCTATCTGCCCTCTTGTCTGTAGAGCTGAAGAGCACCGTC 2630
 Db 2521 CTGCCCCACACGACCGCTATCTGCCCTCTTGTCTGTAGAGCTGAAGAGCACCGTC 2580
 Qy 2631 CCCAGGAGCAGCTCAGGATAGTGTAGTGTGAGCTGTGCGAGGCTTGGCTCCCAAT 2690
 Db 2581 CCCAGGAGCAGCTCAGGATAGTGTGTATGGAGTGTGCCAAGGCTTGGCTCCCAAT 2640
 Qy 2691 AAGCACTAGTCTATAGTGTCTTAGGACTGTGCTGTCACAGCCCGGCGGAGG 2750
 Db 2641 AAGCACTAGTCTATAGTGTCTTAGGACTGTGCTGTCACAGCCCGGCGGAGG 2700
 Qy 2751 CTGCCACACGGAAGCAGCAGATGAATTAATTTCAAGTCAGTGTTTTAAAGAAAGTC 2810
 Db 2701 CTGCCACACGGAAGCAGCAGATGAATTAATTTCAAGTCAGTGTTTTAAAGAAAGTC 2760
 Qy 2811 TTGAAAACAGACCGCGCACCTTTCTCTTAATCCAGCAAAAGTGTATCCCTGTCACACAGA 2870
 Db 2761 TTGAAAACAGACCGCGCACCTTTCTCTTAATCCAGCAAAAGTGTATCCCTGTCACACAGA 2820
 Qy 2871 GACNAGCAGAGTAAACAGGATCAGTGGTCTTAAGTGTCCGAGACTTAACGAAATAGTATT 2930
 Db 2821 GACNAGCAGAGTAAACAGGATCAGTGGTCTTAAGTGTCCGAGACTTAACGAAATAGTATT 2880
 Qy 2931 TCAGCTCAATAAAGATTGAGTTTGCAA 2958
 Db 2881 TCAGCTCAATAAAGATTGAGTTTGCAA 2908

RESULT 3

US-09-564-805-225
 ; Sequence 225, Application US/09564805
 ; Patent No. 6333403
 ; GENERAL INFORMATION:
 ; APPLICANT: Tavtigan, Sean V.
 ; APPLICANT: Teng, David H.F.
 ; APPLICANT: Simard, Jacques
 ; APPLICANT: Rommens, Johanna M.
 ; APPLICANT: Myriad Genetics, Inc.
 ; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
 ; FILE REFERENCE: 2318-258
 ; CURRENT APPLICATION NUMBER: US/09/564,805
 ; PRIOR FILING DATE: 2000-05-05
 ; PRIOR APPLICATION NUMBER: US 60/107,468
 ; PRIOR FILING DATE: 1998-11-06
 ; PRIOR APPLICATION NUMBER: 09/434,382
 ; PRIOR FILING DATE: 1999-11-05
 ; NUMBER OF SEQ ID NOS: 240
 ; SOFTWARE: PatentIn Ver. 2.0

Db 1981 GTGGTCTATTTCGGGGACACATGCGCTCGAGGCTCTGGTCCGATGGGAAGATGCC 2040
Qy 2091 ACCCTCTGTATACATGAAGCCACCTCGAAGATGTTTGGAGAGGAAGCAGTGGAAAG 2150
Db 2041 ACCCTCTGTATACATGAAGCCACCTCGAAGATGTTTGGAGAGGAAGCAGTGGAAAG 2100
Qy 2151 ACACACACACAACTCCCAAGCCATCAGCTGGGGATGCGGATGAACCGGAGTTCAAT 2210
Db 2101 ACACACACACAACTCCCAAGCCATCAGCTGGGGATGCGGATGAACCGGAGTTCAAT 2160
Qy 2211 ATGCTGAACCACTTCAGCAGCGCTATCCCAAGGTCCTCTTCAGCCCACTTCAGC 2270
Db 2161 ATGCTGAACCACTTCAGCAGCGCTATCCCAAGGTCCTCTTCAGCCCACTTCAGC 2220
Qy 2271 GAGAAAGTGGAGTTGCTTTGACACATGAAGTCTCTTTGGAGACTTTTCCAAACAATG 2330
Db 2221 GAGAAAGTGGAGTTGCTTTGACACATGAAGTCTCTTTGGAGACTTTTCCAAACAATG 2280
Qy 2331 CCCAGCTGATTCCTCCCACTGAAAGCCCTGTTTCTGGGACATCGAGAGATGGAGAG 2390
Db 2281 CCCAGCTGATTCCTCCCACTGAAAGCCCTGTTTCTGGGACATCGAGAGATGGAGAG 2340
Qy 2391 CGCAGGAGAAAGCGGAGCTGCGGAGAGTGGGGCGGCCCTCTCTGTCAGGAGCTGGCA 2450
Db 2341 CGCAGGAGAAAGCGGAGCTGCGGAGAGTGGGGCGGCCCTCTCTGTCAGGAGCTGGCA 2400
Qy 2451 GCGGCTCTGGAGATGGGGAGCTTCAGCAGAAAGCGGGGCCCAACAGAGAGCCACAGGCC 2510
Db 2401 GCGGCTCTGGAGATGGGGAGCTTCAGCAGAAAGCGGGGCCCAACAGAGAGCCACAGGCC 2460
Qy 2511 AAGAAGTTCAGAGCCAGTGAAGATCTGGGAGACCTGAACTCAGAGGCTGTGTCTT 2570
Db 2461 AAGAAGTTCAGAGCCAGTGAAGATCTGGGAGACCTGAAATTCAGAGGCTGTGTCTT 2520
Qy 2571 CTGCCCCACGACGACCGCTATCTGCCCTCTGCTGCTGTAAGAGCTGAAGAGCAGCTC 2630
Db 2521 CTGCCCCACGACGACCGCTATCTGCCCTCTGCTGCTGTAAGAGCTGAAGAGCAGCTC 2580
Qy 2631 CCCAGGAGGAGCTCAGGATAGTGGTATGAGAGCTGTGCGAGGCTTTGGGCTCCCAAT 2690
Db 2581 CCCAGGAGGAGCTCAGGATAGTGGTATGAGAGCTGTGCGAGCTGTGCGAGGCTTTAGGCTCCCAAT 2640
Qy 2691 AAGCACTAGTCTATAGATGCTCTTAGACTGCTGCTGGACAGCCGCGGGCCAGGAG 2750
Db 2641 AAGCACTAGTCTATA- -----GGTGCCTGGCACAGCCGCGGGACAGGAG 2684
Qy 2751 CTGCCACACGGAAGCAAGCAGATGAATTAATTTCAATTTCAAGGCGAGTTTAAAGAGTC 2810
Db 2685 CTGCCACACGGAAGCAAGCAGATGAATTAATTTCAATTTCAAGGCGAGTTTAAAGAGTC 2744
Qy 2811 TTGAAAACAGACGGCGGACCTTTCCTCTAATCCAGAAAAGTATTCCTCGACACCCAGA 2870
Db 2745 TTGAAAACAGACGGCAGCACCTTTCCTCTAATCCAGAAAAGTATTCCTCGACACCCAGA 2804
Qy 2871 GACACGAGATTAACAGATCAGTGGGTCTAAGTCTCGAGACTTAAGCAAAATAGTATT 2930
Db 2805 GACACGAGATTAACAGATCAGTGGGTCTAAGTCTCGAGACTTAAGCAAAATAGTATT 2864
Qy 2931 TCAGCTGCAATAAAGATTGAGTTTGCAC 2958
Db 2865 TCAGCTGCAATAAAGATTGAGTTTGCAC 2892

RESULT 4
US-09-564-805-1
; Sequence 1, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtighian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.

; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2481
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
US-09-564-805-1
Query Match 83.9%; Score 2481; DB 4; Length 2481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 51 ATGTGGCGCTTTGCTCGCTGCTGCGGTCCGCGCGCGGACGCCACCATGTCCGAGGAGCG 110
Db 1 ATGTGGCGCTTTGCTCGCTGCTGCGGTCCGCGCGCGGACGCCACCATGTCCGAGGAGCG 60
Qy 111 ACCATATCGCAGGACCCCGCGCGCGGAGCGCGCGCAAGGACCCGCTGCCGACCTG 170
Db 61 ACCATATCGCAGGACCCCGCGCGCGGAGCGCGCGCAAGGACCCGCTGCCGACCTG 120
Qy 171 CGCACGCGAGAGAGCGCGACCGTCCGGGTGCTCCGGCGCGCCAAACACCGTGTACCTG 230
Db 121 CGCACGCGAGAGAGCGCGACCGTCCGGGTGCTCCGGCGCGCCAAACACCGTGTACCTG 180
Qy 231 CAGGTGTGCGCGGGTAGCCGGGACTCCGGGCGCGCGCTCTACGTCCTTCGAGTTTC 290
Db 181 CAGGTGTGCGCGGGTAGCCGGGACTCCGGGCGCGCGCTCTACGTCCTTCGAGTTTC 240
Qy 291 AACCGGTATCTCTTCAACTGTGGAGAGGCGTTTCAGAGACTCATGACAGGACCAAGTTA 350
Db 241 AACCGGTATCTCTTCAACTGTGGAGAGGCGTTTCAGAGACTCATGACAGGACCAAGTTA 300
Qy 351 AAGTTTCTCGCTCGGCAACATATTTCTGCACAGAAATGCACTGGTCTAATGTTGGGGGC 410
Db 301 AAGTTTCTCGCTCGGCAACATATTTCTGCACAGAAATGCACTGGTCTAATGTTGGGGGC 360
Qy 411 TTAAGTGAATGATTTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGA 470
Db 361 TTAAGTGAATGATTTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGA 420
Qy 471 CCTCCACAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTTGAAAGGA 530
Db 421 CCTCCACAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTTGAAAGGA 480
Qy 531 ATAGAACTGCTGTGGGGCCCCACTCTGCCCCAGAAATACGAGGATGAACCATGACAGTT 590
Db 481 ATAGAACTGCTGTGGGGCCCCACTCTGCCCCAGAAATACGAGGATGAACCATGACAGTT 540
Qy 591 TACCAGATCCCCATACACAGTGAACAGAGGGGGAAGGACCAACCATGGCAGAGTTCCA 650
Db 541 TACCAGATCCCCATACACAGTGAACAGAGGGGGAAGGACCAACCATGGCAGAGTTCCA 600
Qy 651 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCCGAATGAAAT 710
Db 601 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCCGAATGAAAT 660
Qy 711 GAGCCACACTTTCACATGTTGTAGCCAGAGAGGGGTTCAGGGACTCTTTCCTGCTGTC 770
Db 661 GAGCCACACTTTCACATGTTGTAGCCAGAGAGGGGTTCAGGGACTCTTTCCTGCTGTC 720

Qy	771	GTAGCTTTTCATCTGTGAAGCTTCACCTTAAAGAGAGGAAACCTTCTGGTGCTCCAAAGCAAAG	830
Db	721	GTAGCTTTTCATCTGTGAAGCTTCACCTTAAAGAGAGGAAACCTTCTGGTGCTCCAAAGCAAAG	780
Qy	831	GAGATGGCCTCCAGCTTGGGACAGCTGCCATCGCTCCCATCATTTGCTGCTCAAGGAC	890
Db	781	GAGATGGCCTCCAGCTTGGGACAGCTGCCATCGCTCCCATCATTTGCTGCTCAAGGAC	840
Qy	891	GGGAAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAAGAGCTGTGTACTCCTCCA	950
Db	841	GGGAAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAAGAGCTGTGTACTCCTCCA	900
Qy	951	GATCCTGGTGCTCTTTTGTGGTAGAATGTCCAGATGAAAGCTTCATTTCAACCCATC	1010
Db	901	GATCCTGGTGCTCTTTTGTGGTAGAATGTCCAGATGAAAGCTTCATTTCAACCCATC	960
Qy	1011	TGTCAGATGCCACCTTTCAGAGGTACCAAGAAAGGCAGATGCCCGTGCCCTTGGTG	1070
Db	961	TGTCAGATGCCACCTTTCAGAGGTACCAAGAAAGGCAGATGCCCGTGCCCTTGGTG	1020
Qy	1071	GTTCCACATGSCCCAGCATCTGTGCTTGTGGACAGCAGGTACACAGCTGGATGGAGAGG	1130
Db	1021	GTTCCACATGSCCCAGCATCTGTGCTTGTGGACAGCAGGTACACAGCTGGATGGAGAGG	1080
Qy	1131	TTTGGGCTTGACACCCAGCACTTGGTGCTCGAATAGAACTGTGCTCAGTTTCAACACTT	1190
Db	1081	TTTGGGCTTGACACCCAGCACTTGGTGCTCGAATAGAACTGTGCTCAGTTTCAACACTT	1140
Qy	1191	CGCAGCACAGATTCNAACCCAGCTCAACCTTCATCCACCCGGACATCTTCCCTGCTC	1250
Db	1141	CGCAGCACAGATTCNAACCCAGCTCAACCTTCATCCACCCGGACATCTTCCCTGCTC	1200
Qy	1251	ACCAGTTTCCGCTGTAGAAGAGAGGSCCCACCCCTCAGTGTGCCATGTTTCAGGGTAA	1310
Db	1201	ACCAGTTTCCGCTGTAGAAGAGAGGSCCCACCCCTCAGTGTGCCATGTTTCAGGGTAA	1260
Qy	1311	TGCTCTCTCAAGTACCAAGCTCCGTCACGAGGAGGTGGCAGAGGATGCCATTTATTACT	1370
Db	1261	TGCTCTCTCAAGTACCAAGCTCCGTCACGAGGAGGTGGCAGAGGATGCCATTTATTACT	1320
Qy	1371	TGCAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCACTTCCACGACAGGCTG	1430
Db	1321	TGCAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCACTTCCACGACAGGCTG	1380
Qy	1431	CAGGAGTACAGGAGGAGTGGCAGGACGSCCCAGCCCCAGCAGAGAGAAAGATCAGTAC	1490
Db	1381	CAGGAGTACAGGAGGAGTGGCAGGACGSCCCAGCCCCAGCAGAGAGAAAGATCAGTAC	1440
Qy	1491	CCAGAAATCATCTTCTCTTGAAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCAGT	1550
Db	1441	CCAGAAATCATCTTCTCTTGAAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCAGT	1500
Qy	1551	GCCACATTTGATCAACATAAGCCCCGACAGTCTCTGTCTACTGGACTGTGGTGGGGCACA	1610
Db	1501	GCCACATTTGATCAACATAAGCCCCGACAGTCTCTGTCTACTGGACTGTGGTGGGGCACA	1560
Qy	1611	TTTGGGCGAGCTGTGCCGTCAATACGAGACACAGGTGGACAGGCTCTGGGCAACCTCGCT	1670
Db	1561	TTTGGGCGAGCTGTGCCGTCAATACGAGACACAGGTGGACAGGCTCTGGGCAACCTCGCT	1620
Qy	1671	GCTGTGTTGTGTCCTCCACCTGACGAGATCACCAACCGGGCTGCCAAGTATCTTCTGTG	1730
Db	1621	GCTGTGTTGTGTCCTCCACCTGACGAGATCACCAACCGGGCTGCCAAGTATCTTCTGTG	1680
Qy	1731	CAGAGAGAACGGGCTTGGCATCTTTGGAAAGGCGCTTACCTTTGCTGGTGGTGGCTGCC	1790
Db	1681	CAGAGAGAACGGGCTTGGCATCTTTGGAAAGGCGCTTACCTTTGCTGGTGGTGGCTGCC	1740
Qy	1791	CCCAACAGAGCTCAAGGCTTGGCTTCCAGCAGTACCAACCCAGTGCAGGAGGCTCTGCAC	1850
Db	1741	CCCAACAGAGCTCAAGGCTTGGCTTCCAGCAGTACCAACCCAGTGCAGGAGGCTCTGCAC	1800
Qy	1851	CACATCAGTATGATTCCTGTCGAAATGCCTTTCAGGAAGGGGCTGAGATCTCCAGTCTGCA	1910

Db	1801	CACATCAGTATGATCTCTGCGCAAAATGCGCTTCAGGAAGGGGCTGAGATCTCCAGTCTCTGCA	1860
Qy	1911	GTGGAAGAGATTGATCAGTTTCGCTGTTTGCAGCAATGTGATTTGGAAAGAGTTTCAGACCTGT	1970
Db	1861	GTGGAAGAGATTGATCAGTTTCGCTGTTTGCAGCAATGTGATTTGGAAAGAGTTTCAGACCTGT	1920
Qy	1971	CTGTGCGGCACCTGCAAGCATGCGTTTGGCTCTGCGCTGGTGCACACCTCTGGCTGGAAA	2030
Db	1921	CTGTGCGGCACCTGCAAGCATGCGTTTGGCTCTGCGCTGGTGCACACCTCTGGCTGGAAA	1980
Qy	2031	GTGTGCTATTTCGGGACACCATGCGCTCGCAGGCTCTGGTCCGATGGGAAAAGATGCC	2090
Db	1981	GTGTGCTATTTCGGGACACCATGCGCTCGCAGGCTCTGGTCCGATGGGAAAAGATGCC	2040
Qy	2091	ACCTCTCTGTATCATGAAGCCACCTTGGNAGATGGTTTGGAAAGGAAGCAGTGGAAAAG	2150
Db	2041	ACCTCTCTGTATCATGAAGCCACCTTGGNAGATGGTTTGGAAAGGAAGCAGTGGAAAAG	2100
Qy	2151	ACACACAGCAACAGTCCCAAGCCATCAGCGTGGGGATGCGGATGAACGCGAGCTTCATT	2210
Db	2101	ACACACAGCAACAGTCCCAAGCCATCAGCGTGGGGATGCGGATGAACGCGAGCTTCATT	2160
Qy	2211	ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGTCTCCCTCTTCAGCCCCAAGTTCAGC	2270
Db	2161	ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGTCTCCCTCTTCAGCCCCAAGTTCAGC	2220
Qy	2271	GAGAAAGTGGGAGTTGCCCTTGACCACATGAAGTCTGCTTTGGAGACTTTCACCAACATG	2330
Db	2221	GAGAAAGTGGGAGTTGCCCTTGACCACATGAAGTCTGCTTTGGAGACTTTCACCAACATG	2280
Qy	2331	CCCAAGCTGATTCCTCCCACTGAAAAGCCCTGTTTGTCTGCGCAGCATCGAGGAGATGGAGAG	2390
Db	2281	CCCAAGCTGATTCCTCCCACTGAAAAGCCCTGTTTGTCTGCGCAGCATCGAGGAGATGGAGAG	2340
Qy	2391	CGCAGGAGAAAGCGGAGTTCGGCAGGTGCGGCGCGGCTCAGCAGAGCGGGCCCTCTCTTCAGGGAGCTCGCA	2450
Db	2341	CGCAGGAGAAAGCGGAGTTCGGCAGGTGCGGCGCGGCTCAGCAGAGCGGGCCCTCTCTTCAGGGAGCTCGCA	2400
Qy	2451	GGCGGCTCGAGATGGGAGCTCAGCAGAGCGGGCCCTCTCTTCAGGGAGCTCGGCA	2510
Db	2401	GGCGGCTCGAGATGGGAGCTCAGCAGAGCGGGCCCTCTCTTCAGGGAGCTCGGCA	2460
Qy	2511	AAGAAAGTCAAGCCCCAGTGA	2531
Db	2461	AAGAAAGTCAAGCCCCAGTGA	2481
RESULT 5			
US-09-564-805-221			
; Sequence 221, Application US/09564805			
; Patent No. 6333403			
; GENERAL INFORMATION:			
; APPLICANT: Tavtigian, Sean W.			
; APPLICANT: Teng, David H.P.			
; APPLICANT: Simard, Jacques			
; APPLICANT: Rommens, Johanna M.			
; APPLICANT: Myriad Genetics, Inc.			
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility			
; FILE OF INVENTION: Gene and a Palalog and Orthologous Genes			
; FILE REFERENCE: 2318-258			
; CURRENT APPLICATION NUMBER: US/09/564,805			
; CURRENT FILING DATE: 2000-05-05			
; PRIOR APPLICATION NUMBER: US 60/107,468			
; PRIOR FILING DATE: 1998-11-06			
; PRIOR APPLICATION NUMBER: 09/434,382			
; PRIOR FILING DATE: 1999-11-05			
; NUMBER OF SEQ ID NOS: 240			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 221			
; LENGTH: 2470			
; TYPE: DNA			
; ORGANISM: Mus musculus			

RESULTS

```

RES001 5
US-09-564-805-221
; Sequence 221, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtighian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linker
; TITLE OF INVENTION: Gene and a Paralog a
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-03-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 221
; LENGTH: 2470
; TYPE: DNA
; ORGANISM: Mus musculus

```


FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2466)
US-09-564-805-221

Query Match 55.6%; Score 1645.6; DB 4; Length 2470;

Best Local Similarity 81.6%; Pred. No. 0;

Matches 1958; Conservative 0; Mismatches 417; Indels 24; Gaps 4;

```
Qy 108 GCGACCATATCGCAGGACACCGCCGCGAGCGCGCCGCAAGACCCGCTGCGGCAC 167
Db 40 GCGACCATGTGCGAGGTTGCGCTGCTGCGCGCGCCGCGCCGCAAGACCCGCTGCGGCAC 99
Qy 168 CTGCGCAGCGAGAGAGCGGACCGCTGCGGGTCTCTCGGGGCCCAACACCGTGTAC 227
Db 100 CTGCGTACCGGGAGAGCGGGCC-----GGGTCCCGGGGCCGGAACCGCTGTAC 153
Qy 228 CTGCAAGGTGGTGGCAGCGGGTAGCCGGACTCGGGCGCGCGCTCTACGTCTCTTCCGAG 287
Db 154 CTGCAAGGTGGTGGCGCGGGCGCGCGCGGAGCGGGGGGCTGCTCTATGTCTTCTCGGAA 213
Qy 288 TTCAACCGGTATCTCTTCAACTGTGAGAAGCGGTTCTAGAGACTCATCGAGGACACAAG 347
Db 214 TACAACAGGTACTTTTAACTGCGGAGAGCGGCTCCAAAGACTTATGACGGAACACAAG 273
Qy 348 TTAAGGTGTGCGCTGGACACATATTTCTGACACAAATGCACTGTCTTAATGTTGGG 407
Db 274 ACTGAAGTGTGCTGCTTGACAACATCTTCTGACTCGGATGCAATGTTGGG 333
Qy 408 GCGTTAAGTGGAAATGATTTTAAAGGAAACCGGGCTTCAAAGTGTGTACTTTCT 467
Db 334 GGGTTGTGGNAATTTTAACTTTAAAGGAAACCGGGCTTCCAAATGTTCTGTCT 393
Qy 468 GGACCTCCACAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTTGAA 527
Db 394 GGACCACACAGCTGGAGAAATATCTAGAAGCAATCAAAATATTTTCTGGTCCATTTGAA 453
Qy 528 GGAATAGAACTGGCTGGCGGCCCACTCTGCCAGAAATACGAGGATGAACCAATGACA 587
Db 454 GGAATAGAACTGGCGTGGCGCTCACTCTGCACCAAGATCAAGGATGAGACCATGACT 513
Qy 588 GTTTACCAAGTCCCATACAGTGAACAGAGAGGGGAAAGCACCAACATGGCAGAGT 647
Db 514 GTTTACCAAGTCCCATACAGTGAACAGAGGAGTGTGAAGCAACAGCCATCCAGAGC 573
Qy 648 CCAGAAAGCGCTCTCAGCAGGCTCAGTCCAGAGCATCTTCCAGATCCCGAGTCGAATGAA 707
Db 574 CCCAGAACATCTCCCAACAGGCTCAGTCCCAACAGTCAATCGGACTCTGGATCAGCTGAA 633
Qy 708 AATGAGCCACACCTTCCACATGTTGTAGCCAGAGAGAGGGGTCAAGGACTCTTCCCTG 767
Db 634 AATGGGC-----AGTCCCAACAGGAAGCATGGGCGAGGAC-CCTCTCTTA 678
Qy 768 GTCGTAGCTTTTCATCTGAAGCTTCACTTAAAGAGAGAACTCTTGGTGTCAAAGCA 827
Db 679 GTGGTAGCTTTTGTCTGAAGCTTCACTTGAAGAAAGAACTCTTGGTGTCAAAGCA 738
Qy 828 AAGGAGATGGGCTCCAGTGTGGGACAGTGCATCGCTCCCATCAATGCTGCTGTCAAG 887
Db 739 AAGGAGTGGGCTTCTTGTGGGACGCGCCCATTTGCAACCATCAATGCTGTCAAG 798
Qy 888 GACGGAAAGCATCACTCATGAAGAGAGAGATTTGGCTGAAGAGCTGTGTACTCT 947
Db 799 GACGGAAAGATATCACTTACGAAGGAAGAGATTTGCTGTGAAGAGCTTTGTACACC 858
Qy 948 CCAGATCTGTGTGCTTTTGTGTGTGTAGAACTGTCCAGATGAAGCTTCATTCACACC 1007
Db 859 CCAGATCTGTGTGCTTTTGTGTGTGTAGTGTCTGTGTGAAGAGATTCATCTGCGCC 918
Qy 1008 ATCTGTGAAGATGCCACCTTTTCAGAGGTACCAAGGAAAGGAGGATGCCCGCTGGCTTG 1067
Db 919 ATCTGTGAGACGACACCTTTTAAAGGTACCAAGGAGAGGCTGTATGCACCTGTGGCGTG 978
```

```
Qy 1068 GTGTTTCCACATGCCCCCAGCATCTGTCTTGTGACAGCAGGTACACAGCAGTGGATCGAG 1127
Db 979 GTGTTCCACATAGCCCCCAGAACTCTGTACTCATCGACAGCAGATACACAGCAGTGGATCGAG 1038
Qy 1128 AGGTTTGGGCTTGACACCCAGCAGCTTGGTCTCTGAATGAGAAGTGTGCTTCAGTTCAAC 1187
Db 1039 AGGTTGGGCTTGACACACAGCAGCTGATTTCTGAATGAGAATTTGCCCTCGTCCACAAC 1098
Qy 1188 CTTCGAGCAGCAGAGATTCAAAACCCAGCTCAACTCATCCACCGGACATCTTCCCCCTG 1247
Db 1099 CTGCGAGCCACAGATTTCAGAGCCAGCTCAGCTCATCCACCTGACATCTTCCCCCTG 1158
Qy 1248 CTACCAAGTTTTCGCTGTAGAGAGGGGCCCCACCTCAGTGTGCCATGTTTCAGGGT 1307
Db 1159 CTTACCAAGCTTCTATAGTAAGGAGGAGGTCCACCTCAGCTGCGCAACAGTTTCGGGGT 1218
Qy 1308 GAATGCTCTCTCAAGTACCAAGCTCCGTCACGAGGAGGTGGCAGAGGATGCCATTAAT 1367
Db 1219 GAATGCTCTCTCAAGTATTCAGTCCGCCCCAAGAGAGAGTGGCAGAGGATACCACTC 1278
Qy 1368 ACTTGCAATCTCAGGAAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCCAGCAGAGC 1427
Db 1279 GACTGCAATCTCATGATGAATTCATAGCTGAGGCTTGGAGCTCCCGAGTTTCCAGGAGT 1338
Qy 1428 GTCAGAGGTACAGGAGGTGCGCAGGACGGCCCGCCAGCCAGCAGAGAGAAAGAGTCAG 1487
Db 1339 GTGAGAGGTATTCGGAAGAAACGTGCGAGGAAACCCAGCCCCAGCAGAGAGAAAGAGCCAG 1398
Qy 1488 TACCCAGAAATCATCTTCTTGGAAACAGGCTCTGCCATCCCGATGAAAGTTCGAAATGTC 1547
Db 1399 TATCTGAAATTCCTTCTGGGTACGGGTCTGCCATCCCAATGGAGATCCGAAATGTC 1458
Qy 1548 AGTGCCACATCTGTCAACATAAGCCCGCACAGCTCTCTGTCTCTGTAATGAGTGTGAGGGC 1607
Db 1459 AGTTCCACATCTCGTCAACCTTAAGCCCTGACAAGTCACTGCTCTGATTTGGAGAGAGC 1518
Qy 1608 ACATTTGGGCGAGTGTGCGCTCATACGGAGACCGGTGACAGGCTCTGGGACCCCTG 1667
Db 1519 ACTTTGGGCGAGTGTGCGCTCATACGGAGACCAATAGACCGAGTCTTATGCGGCTC 1578
Qy 1668 GCTGCTGTGTTTGTCTCCACCTCGACGAGATCACACACGGGCTTGGCAAGTATCTTG 1727
Db 1579 ACGGCTGTGTTGTCTCCACCTCGACCGCCAGCCACACAGGCTTGTGTAATATCTTG 1638
Qy 1728 CTGAGAGAGAAACGCGCTTGGCATCTTTGGGAAAGCCGCTTCACTTCTGCTGTGTTGGT 1787
Db 1639 CTGAGAGAGAGCATGCGTTGGCATCTCTGGGAAACCTTCCAGCCCTTGTGTTGGTG 1698
Qy 1788 GCGCCCAACAGCTCAAGCCTGGCTCCAGCAGTACCACACAGTCCAGGAGTCCCTG 1847
Db 1699 GCTCTTACCCAGCTCAGGGCTGGCTGCGCAGTATCACAACTCCAGGAGATTCG 1758
Qy 1848 CACCAATCATGATGATTCCTGCAAAATGCTTTCAGGAAGGGCTGAGATCTCCAGTCT 1907
Db 1759 CACCACTCATGATGATTCCTGCAAAATGCTTTCAGAAAGGGCAGAGTCTCCAATACT 1818
Qy 1908 GCAGTGAAGAGATTGATCAGTTCGCTTGTGCGAAACATGTGATTTGGAAGAGTTTCAGACC 1967
Db 1819 ACATTGAAAGGCTGATAAGCTTGTGTTGGAAACATGTGACTTAGAAGAAATTTCAAGC 1878
Qy 1968 TGTCTGTGCGGACATGCAAGCATGTTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2027
Db 1879 TGCCTGTGTAAGGACATGCAAGCATGCTTTTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1938
Qy 2028 AAAGTGTCTTATTTCCGGGACACCATGCCCTGGAGGCTCTGCTCCGATGCGGAAAGAT 2087
Db 1939 AAGTCTCTACTCTGGGGATACCATGCTCTGTGAGCTCTGCTGCTGCTGCTGCTGCTGCTG 1998
Qy 2088 GCCACCTCTCTGATACATGAAGCCACCTCGAAGATGGTGTGGAAGAGGAGCAGTGGAA 2147
Db 1999 GCCACCTCTCTGATACATGAAGCCACCTCTGAGGATCNCCTTGGAAAGGAGGAGCAGTAGAG 2058
Qy 2148 AAGACACACAGCACAACGTCCCAAGCCATCAGCGTGGGATGCGGATGAACGCGGAGTTC 2207
```


Db 2059 AGACACACAGACACACCTCCAGGCTATTAATGTGGGATCGGATGAATCGGAGTTC 2118
QY 2208 ATTATGCTGAACACACTTCAGCCAGCCTATGCAAGGTCCCCCTCTTTCAGCCCCCACTTC 2267
Db 2119 ATCATGCTGAACACACTTCAGTCAGCGGTACGACGATCCCTCTTTCAGCCCTGACTTC 2178
QY 2268 AGCGAAGAGTGGGAGTTCCTTTGACCAACATGAAGGTCTGCTTTGGAGACTTTTCAACA 2327
Db 2179 AAGCGAAGAGTGGGAGTTCCTTTGACCAACATGAAGGTCTGNTTTGGAGACTTCCGACA 2238
QY 2328 ATGCCCAAGCTGATTTCCCACTGAAGCCCTGTTTGTGGCGACATCAGAGATGGAG 2387
Db 2239 GTGCCCAAGCTGATTTCCCACTGAAGGCCCTGTTTGTGGCGACATCAGAGATGGAG 2298
QY 2388 GAGCGCAGGGAAGAGGAGGTGCGGCGAGGTGCGGCGGCCCTCTGTCAGGAGAGTGG 2447
Db 2299 GAACGAGGGAAGAGGAGGTACGGCTGTGCGAGAGCCCTCTGACC---CAGCAG 2355
QY 2448 GCAGGCGGCTGGAGATGGGAGCTTCAGCAGAACGCGGCGCCACACAGAGAGCCACA 2506
Db 2356 GCAGACAGCCAGAGGACAGAGAACCCCAACAGAGAGCGGCGCCACACAGATGAACACA 2414

RESULT 6

US-09-564-805-28
; Sequence 28, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigan, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 26664
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (910)..(13104)
; OTHER INFORMATION: exon 1: 910-1154; exon 2: 1736-1786; exon 3:
; OTHER INFORMATION: 1925-1995; exon 4: 3025-3089; exon 5: 4361-4418;
; OTHER INFORMATION: exon 6: 5582-5650; exon 7: 7075-7194; exon 8:
; OTHER INFORMATION: 8186-8244; exon 9: 12878-12936; exon 10:
; NAME/KEY: misc feature
; LOCATION: (13756)..(22917)
; OTHER INFORMATION: exon 11: 13756-13868; exon 12: 15283-15378; exon
; OTHER INFORMATION: 13: 16278-16416; exon 14: 16498-16583; exon 15:
; OTHER INFORMATION: 18583-18701; exon 16: 20349-20445; exon 17:
; OTHER INFORMATION: 22172-22310; exon 18: 22879-22917
; NAME/KEY: misc feature
; LOCATION: (23045)..(26452)
; OTHER INFORMATION: exon 19: 23045-23154; exon 20: 23795-23895; exon
; OTHER INFORMATION: 21: 23973-24093; exon 22: 24354-24432; exon 23:
; OTHER INFORMATION: 25026-25170; exon 24: 25812-26036; polyadenylation
; OTHER INFORMATION: signal: 26447-26452
; NAME/KEY: variation
; LOCATION: (826)..(23879)
; OTHER INFORMATION: s at positions 826 and 23180 is G or C; y at
; OTHER INFORMATION: positions 1914, 5568, 7165, 16431, 1857 and 20486
; OTHER INFORMATION: is C or T; n at position 13128 is t or tgat; r at

; OTHER INFORMATION: positions 22211 and 23879 is A or G.
US-09-564-805-28

Query Match 22.2%; Score 657.2; DB 4; Length 26664;
Best Local Similarity 99.5%; Pred. No. 9.6e-167;
Matches 659; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2297 CATGAAGTCTGCTTTGGAGACTTTTCAACAATGCCAAGCTGATTTCCCCCTCTGAAAGC 2356
Db 25805 CTTCTAGGTCTGCTTTGGAGACTTTTCAACAATGCCAAGCTGATTTCCCCCTCTGAAAGC 25864
QY 2357 CTTGTTTCTGCTGGCAGATCGAGGAGATGGAGAGCGCAGGAGAGCGGAGCTCGGCA 2416
Db 25865 CTTGTTTCTGCTGGCAGATCGAGGAGATGGAGAGCGCAGGAGAGCGGAGCTCGGCA 25924
QY 2417 GGTGCGGCGGCCCTCTGTCAGGAGCTGGCAGCGGCCCTGGAGGATGGGAGCTCTCA 2476
Db 25925 GGTGCGGCGGCCCTCTGTCAGGAGCTGGCAGCGGCCCTGGAGGATGGGAGCTCTCA 25984
QY 2477 GCAGAAGCGGCGCCACACAGAGGAGCCACAGGCCAAGAGGTCAGAGGCCAGTGAAGATC 2536
Db 25985 GCAGAAGCGGCGCCACACAGAGGAGCCACAGGCCAAGAGGTCAGAGGCCAGTGAAGATC 26044
QY 2537 TGGGAGACCTGAACTCAGAAGGCTGTGTCTTCTGCCCCACGACGACCCGCTATCTG 2596
Db 26045 TGGGAGACCTGAACTCAGAAGGCTGTGTCTTCTGCCCCACGACGACCCGCTATCTG 26104
QY 2597 CCCTCTTGTGTTAGAGCTGAAGAGCAGCGTCCCCCAGGAGGAGCTCAGGATAGGTG 2656
Db 26105 CCCTCTTGTGTTAGAGCTGAAGAGCAGCGTCCCCCAGGAGGAGCTCAGGATAGGTG 26164
QY 2657 GTATGAGCTGTGCGGAGGCTTGGGCTCCACATGAAGCACTAGTCTATAGATGCTCTT 2716
Db 26165 GTATGAGCTGTGCGGAGGCTTGGGCTCCACATGAAGCACTAGTCTATAGATGCTCTT 26224
QY 2717 GGACTGTGCTTGGCAGCAGCGCGGCCAGGAGGCTGCCACGAGAGCAGCATGAA 2776
Db 26225 GGACTGTGCTTGGCAGCAGCGCGGCCAGGAGGCTGCCACGAGAGCAGCATGAA 26284
QY 2777 CTAATTTCAATTTCAAGGAGCTTTTAAAGAGAGTCTTGAACAGACGCGGACCTTTCC 2836
Db 26285 CTAATTTCAATTTCAAGGAGCTTTTAAAGAGAGTCTTGAACAGACGCGGACCTTTCC 26344
QY 2837 TCTAATCCAGCAAGTGTATTCCTGCACACAGAGAGCAGAGTAAAGATGAGTGG 2896
Db 26345 TCTAATCCAGCAAGTGTATTCCTGCACACAGAGAGCAGAGTAAAGATGAGTGG 26404
QY 2897 GTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTTTCAGTGCATTAAGATTTGCT 2956
Db 26405 GTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTTTCAGTGCATTAAGATTTGCT 26464
QY 2957 AA 2958
Db 26465 AA 26466

RESULT 7

US-09-564-805-27
; Sequence 27, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigan, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06

; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 655
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(228)
; OTHER INFORMATION: exon 24
; NAME/KEY: polyA signal
; LOCATION: (636)..(641)
US-09-564-805-27

Query Match 22.1%; Score 655; DB 4; Length 655;
Best Local Similarity 100.0%; Pred. No. 5.2e-167;
Matches 655; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2304 GTCTGCTTTGGAGACTTCCAAATGCGCAAGCTGATTCGCCCACTGAAAGCCCTGTTT 2363
Db 1 GTCTGCTTTGGAGACTTCCAAATGCGCAAGCTGATTCGCCCACTGAAAGCCCTGTTT 60
QY 2364 GCTGCGCATCGAGGAGATGAGGAGCGGAGGAGAGCGGAGCTCGCGCAGTGCGG 2423
Db 61 GCTGCGCATCGAGGAGATGAGGAGCGGAGGAGAGCGGAGCTCGCGCAGTGCGG 120
QY 2424 GCGGCCCTCTGTCAGGAGAGTGGCAGCGGCTTGGAGGATGGGAGCCCTCAGCAGAG 2483
Db 121 GCGGCCCTCTGTCAGGAGAGTGGCAGCGGCTTGGAGGATGGGAGCCCTCAGCAGAG 180
QY 2484 CGGGCCACACAGAGAGCCACAGGCCAAGAGTGCAGAGCCAGTGAAGATCTGGGAGA 2543
Db 181 CGGGCCACACAGAGAGCCACAGGCCAAGAGTGCAGAGCCAGTGAAGATCTGGGAGA 240
QY 2544 CCTCAACTCAGAGGCTGTGTCTCTGCCCCACGACGACCCGCTATGCCCCCTCT 2603
Db 241 CCTCAACTCAGAGGCTGTGTCTCTGCCCCACGACGACCCGCTATGCCCCCTCT 300
QY 2604 TGCTGGTGAAGCTGAAGAGCACCGTCCGCCAGGAGGAGCTCAGGATAGTGGTATGGA 2663
Db 301 TGCTGGTGAAGCTGAAGAGCACCGTCCGCCAGGAGGAGCTCAGGATAGTGGTATGGA 360
QY 2664 GCTGTGCGAGCTGGGCTCCCAATAGACTAGTCTATAGTGCCTCTTAGACTGG 2723
Db 361 GCTGTGCGAGCTTGGGCTCCCAATAGACTAGTCTATAGTGCCTCTTAGACTGG 420
QY 2724 TGCTGGGACACGCGGCGGAGGAGTGCACACGGAAGCAGATGAATTAATTT 2783
Db 421 TGCTGGGACACGCGGCGGAGGAGTGCACACGGAAGCAGATGAATTAATTT 480
QY 2784 CATTTCAAGGCAAGTTTAAAGAGTCTTTGAAACAGACGCGGACCTTTCTCTAATC 2843
Db 481 CATTTCAAGGCAAGTTTAAAGAGTCTTTGAAACAGACGCGGACCTTTCTCTAATC 540
QY 2844 CAGCAAGTGTATCCCTGCACACAGAGACAGAGTAACAGGATCAGTGGGTCTAAG 2903
Db 541 CAGCAAGTGTATCCCTGCACACAGAGACAGAGTAACAGGATCAGTGGGTCTAAG 600
QY 2904 TGTCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAGATTGAGTTTGCAA 2958
Db 601 TGTCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAGATTGAGTTTGCAA 655

RESULT 8
US-09-564-805-210
; Sequence 210, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques

; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 210
; LENGTH: 350
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (51)..(293)
US-09-564-805-210

Query Match 10.1%; Score 297.4; DB 4; Length 350;
Best Local Similarity 98.0%; Pred. No. 9.9e-71;
Matches 301; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGCGGGGTAGGTGACCGGCGGCTTTCTCAGTTTGGTGAGAGCGGCGCATGTGGGCGC 60
Db 1 CGCGGGGTAGGTGACCGGCGGCTTTCTCAGTTTGGTGAGAGCGGCGCATGTGGGCGC 60
QY 61 TTTGCTCGCTGCTGCGGCTCCGCGCGGACGACCATGTCCAGGGAGCGACCATATCGC 120
Db 61 TTTGCTCGCTGCTGCGGCTCCGCGCGGACGACCATGTCCAGGGAGCGACCATATCGC 120
QY 121 AGGCACCGGCGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
Db 121 AGGCACCGGCGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
QY 181 AGAAGCGGAGCGTCCGGGTGCTCCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
Db 181 AGAAGCGGAGCGTCCGGGTGCTCCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
QY 241 CAGCGGGTAGCGGGGACTCCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
Db 241 CAGCGGGTAGCGGGGACTCCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
QY 301 TCTTCAA 307
Db 301 TCAACGA 307

RESULT 9
US-09-564-805-4
; Sequence 4, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4

```
; LENGTH: 295
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51)..(295)
; OTHER INFORMATION: exon 1
US-09-564-805-4

Query Match      10.0%; Score 295; DB 4; Length 295;
Best Local Similarity 100.0%; Pred. No. 4e-70;
Matches 295; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCGGGCGGTAGTACCGCGCGCTTCTCAGTTTGGTGGAGACGGGCCATGTGGGGCC 60
Db 1 CCGGGCGGTAGTACCGCGCGCTTCTCAGTTTGGTGGAGACGGGCCATGTGGGGCC 60

Qy 61 TTTGCTCGCTGCTGGGTCGCGCGCGGACGACCATGTGCGAGGACCGACCATATCGC 120
Db 61 TTTGCTCGCTGCTGGGTCGCGCGCGGACGACCATGTGCGAGGACCGACCATATCGC 120

Qy 121 AGGACCCCGCGCGCGAGCGCGCGCAAGGACCCGCTGCGGACCTGCGCAGCGCGAG 180
Db 121 AGGACCCCGCGCGCGAGCGCGCGCAAGGACCCGCTGCGGACCTGCGCAGCGCGAG 180

Qy 181 AGAAGCGCGGACCGTGGGGTGTCCGGCGGCCCAACACCGTGTACCTGCAAGTGGTG 240
Db 181 AGAAGCGCGGACCGTGGGGTGTCCGGCGGCCCAACACCGTGTACCTGCAAGTGGTG 240

Qy 241 CAGCGGTTAGCGGACTCGGCGCGCGCTCTACGCTCTCTCCGAGTTCAACCG 295
Db 241 CAGCGGTTAGCGGACTCGGCGCGCGCTCTACGCTCTCTCCGAGTTCAACCG 295

RESULT 10
US-09-328-111-315
; Sequence 315, Application US/09328111
; Patent No. 6262333
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Derti, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/328,111
; CURRENT FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: US 60/088,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 315
; LENGTH: 238
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-328-111-315

Query Match      8.0%; Score 237; DB 4; Length 238;
Best Local Similarity 100.0%; Pred. No. 1.6e-54;
Matches 237; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 226 ACCTGAGGTGGTGGCAGCGGTAGCGGGACTCGGGCGCGGCTCTACGCTCTCTCGG 285
Db 1 ACCTGAGGTGGTGGCAGCGGTAGCGGGACTCGGGCGCGGCTCTACGCTCTCTCGG 285

Qy 286 AGTTCAACCGGTATCTTCAACTGTGTGAGAGGCGTTTCAAGACTCATGTGAGGAGCACA 345
Db 61 AGTTCAACCGGTATCTTCAACTGTGTGAGAGGCGTTTCAAGACTCATGTGAGGAGCACA 120

Qy 346 AGTTAAAGGTTGCTCGCCTGGACACATATTCCTTACACAGGAATGCACCTGTTTAAATGTTG 405
Db 121 AGTTAAAGGTTGCTCGCCTGGACACATATTCCTTACACAGGAATGCACCTGTTTAAATGTTG 180

Qy 406 GGGGCTTAAGTGAATGATTTCTTAAAGGAAACCGGGCTTCCAAAGTGTGTAC 462
Db 181 GGGGCTTAAGTGAATGATTTCTTAAAGGAAACCGGGCTTCCAAAGTGTGTAC 237

RESULT 11
US-09-564-805-26
; Sequence 26, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
; LENGTH: 145
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(145)
; OTHER INFORMATION: exon 23
US-09-564-805-26

Query Match      4.9%; Score 145; DB 4; Length 145;
Best Local Similarity 100.0%; Pred. No. 7.7e-30;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2159 CACAACGTCCTCAAGCCATCAGCGTGGGATCGGATGAACGCGAGTTTCATTATGCTGAA 2218
Db 1 CACAACGTCCTCAAGCCATCAGCGTGGGATCGGATGAACGCGAGTTTCATTATGCTGAA 60

Qy 2219 CCATTTCAGCCAGCGCTATGCCAAGTCCCTCTTTCAGCCCCCACTTCAGCGAGAAAGT 2278
Db 61 CCATTTCAGCCAGCGCTATGCCAAGTCCCTCTTTCAGCCCCCACTTCAGCGAGAAAGT 120

Qy 2279 GGGAGTTCCTTGCACCATGAAG 2303
Db 121 GGGAGTTCCTTGCACCATGAAG 145

RESULT 12
US-09-564-805-16
; Sequence 16, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
```

;; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
;; FILE REFERENCE: 2318-258
;; CURRENT APPLICATION NUMBER: US/09/564,805
;; CURRENT FILING DATE: 2000-05-05
;; PRIOR APPLICATION NUMBER: US 60/107,468
;; PRIOR FILING DATE: 1998-11-06
;; PRIOR APPLICATION NUMBER: 09/434,382
;; PRIOR FILING DATE: 1999-11-05
;; NUMBER OF SEQ ID NOS: 240
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 16
;; LENGTH: 139
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (1)..(139)
;; OTHER INFORMATION: exon 13
US-09-564-805-16

Query Match 4.7%; Score 139; DB 4; Length 139;
Best Local Similarity 100.0%; Pred. No. 3.1e-28;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1130 GTTTGGGCTGACACCCAGCACTTGGTCTGTAATGAGAACTGTGCTCAGTTCAACAACCT 1189
Db 1 GTTTGGGCTGACACCCAGCACTTGGTCTGTAATGAGAACTGTGCTCAGTTCAACAACCT 60
Qy 1190 TCGAGCCACAAAGATTCAAAACCCAGCTCAACCTCATCCACCGGACATCTTCCCCCTGCT 1249
Db 61 TCGAGCCACAAAGATTCAAAACCCAGCTCAACCTCATCCACCGGACATCTTCCCCCTGCT 120
Qy 1250 CACCAGTTTCGCTGTAAG 1268
Db 121 CACCAGTTTCGCTGTAAG 139

RESULT 13
US-09-564-805-20
; Sequence 20, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigan, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 139
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(139)
; OTHER INFORMATION: exon 17
US-09-564-805-20

Query Match 4.7%; Score 139; DB 4; Length 139;
Best Local Similarity 100.0%; Pred. No. 3.1e-28;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1571 CCCCAGACGTCCTGCTACTGAGCTGTGGTGGGACACATTGGGCGACCTGTGCCGTCA 1630

Db 1 CCCCAGACGTCCTGCTACTGAGCTGTGGTGGGACACATTGGGCGACCTGTGCCGTCA 60
Qy 1631 TTACGGAGACAGGTGACAGGGTCTTGGGACACCTGGCTGCTGTGTTGTGTCACCT 1690
Db 61 TTACGGAGACAGGTGACAGGGTCTTGGGACACCTGGCTGCTGTGTTGTGTCACCT 120
Qy 1691 GCAGCGAGATCACCACAG 1709
Db 121 GCAGCGAGATCACCACAG 139

RESULT 14
US-09-564-805-24
; Sequence 24, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigan, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 121
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(121)
; OTHER INFORMATION: exon 21
US-09-564-805-24

Query Match 4.1%; Score 121; DB 4; Length 121;
Best Local Similarity 100.0%; Pred. No. 2.1e-23;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1959 TTTCAGACCTCTCTGGTCGGCACTGCAAGCATGCGTTTGGCTGTGCGCTGGTGACACC 2018
Db 1 TTTCAGACCTCTCTGGTCGGCACTGCAAGCATGCGTTTGGCTGTGCGCTGGTGACACC 60
Qy 2019 TCTGGTGGAAAGTGGTCTATTCCGGGGACACCATGCCCTCGAGGCTCTGTCGGGATG 2078
Db 61 TCTGGTGGAAAGTGGTCTATTCCGGGGACACCATGCCCTCGAGGCTCTGTCGGGATG 120
Qy 2079 G 2079
Db 121 G 121

RESULT 15
US-09-564-805-10
; Sequence 10, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigan, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258

; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(120)
; OTHER INFORMATION: exon 7
US-09-564-805-10

Query Match 4.1%; Score 120; DB 4; Length 120;
Best Local Similarity 100.0%; Pred. No. 3.8e-23;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 610 GTCAACAGAGGAGGGGAAGCACCACCATGGCAGAGTCCAGAAAGGCCCTCTCAGCAGGC 669
|||
Db 1 GTCAACAGAGGAGGGGAAGCACCACCATGGCAGAGTCCAGAAAGGCCCTCTCAGCAGGC 60
|||
Oy 670 TCAGTCCAGAGCGATCTTCAGACTCCGAGTCGAATGAAATGAGCCACACCTTCCACATG 729
|||
Db 61 TCAGTCCAGAGCGATCTTCAGACTCCGAGTCGAATGAAATGAGCCACACCTTCCACATG 120
|||

Search completed: May 18, 2003, 00:09:50
Job time : 249 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 18, 2003, 00:09:53 ; Search time 359 Seconds
(without alignments)
10628.516 Million cell updates/sec

Title: US-09-434-382-3
Perfect score: 2958
Sequence: 1 cgcggcgtagtgaccggc.....ataaagattgatttgcga 2958

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 810007 seqs, 644960991 residues

Total number of hits satisfying chosen parameters: 1620014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2958	100.0	2958	9	US-09-988-626-3
2	2958	100.0	2958	9	US-09-988-687-3
3	2874.4	97.2	2908	9	US-09-988-626-223
4	2874.4	97.2	2908	9	US-09-988-687-223
5	2819.6	95.3	2892	9	US-09-988-626-225
6	2819.6	95.3	2892	9	US-09-988-687-225
7	2481	83.9	2481	9	US-09-988-626-1
8	2481	83.9	2481	9	US-09-988-687-1
9	1645.6	55.6	2470	9	US-09-988-626-221
10	1645.6	55.6	2470	9	US-09-988-687-221
11	734.8	24.8	783	10	US-09-833-381-2039
12	657.2	22.2	2664	9	US-09-988-626-28
13	657.2	22.2	2664	9	US-09-988-687-28
14	655	22.1	655	9	US-09-988-626-27
15	655	22.1	655	9	US-09-988-687-27
16	470.4	15.9	536	10	US-09-833-381-2038
17	432.8	14.6	554	9	US-09-918-995-8996
18	297.4	10.1	350	9	US-09-988-626-210
19	297.4	10.1	350	9	US-09-988-687-210

Sequence 4, Appli
Sequence 4, Appli
Sequence 315, App
Sequence 26, Appl
Sequence 26, Appl
Sequence 16, Appl
Sequence 20, Appl
Sequence 20, Appl
Sequence 24, Appl
Sequence 24, Appl
Sequence 24, Appl
Sequence 10, Appl
Sequence 10, Appl
Sequence 18, Appl
Sequence 18, Appl
Sequence 212, App
Sequence 212, App
Sequence 14, Appl
Sequence 14, Appl
Sequence 22, Appl
Sequence 22, Appl
Sequence 23, Appl
Sequence 23, Appl
Sequence 19, Appl
Sequence 19, Appl
Sequence 15, Appl

20 295 10.0 295 9 US-09-988-626-4
21 295 10.0 295 9 US-09-988-687-4
22 237 8.0 238 10 US-09-879-536-315
23 145 4.9 145 9 US-09-988-626-26
24 145 4.9 145 9 US-09-988-687-26
25 139 4.7 139 9 US-09-988-626-16
26 139 4.7 139 9 US-09-988-626-20
27 139 4.7 139 9 US-09-988-687-16
28 139 4.7 139 9 US-09-988-687-20
29 121 4.1 121 9 US-09-988-626-24
30 121 4.1 121 9 US-09-988-687-24
31 120 4.1 120 9 US-09-988-626-10
32 120 4.1 120 9 US-09-988-687-10
33 119 4.0 119 9 US-09-988-626-18
34 119 4.0 119 9 US-09-988-687-18
35 118.4 4.0 326 9 US-09-988-626-212
36 118.4 4.0 326 9 US-09-988-687-212
37 113 3.8 113 9 US-09-988-626-14
38 113 3.8 113 9 US-09-988-687-14
39 110 3.7 110 9 US-09-988-626-22
40 110 3.7 110 9 US-09-988-687-22
41 100 3.4 100 9 US-09-988-626-23
42 100 3.4 100 9 US-09-988-687-23
43 97 3.3 97 9 US-09-988-626-19
44 97 3.3 97 9 US-09-988-687-19
45 96 3.2 96 9 US-09-988-626-15

ALIGNMENTS

RESULT 1
US-09-988-626-3
; Sequence 3, Application US/09988626
; Publication No. US20030044959A1
; GENERAL INFORMATION:
; APPLICANT: Tavtighian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,626
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2958
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51)..(2531)
; OTHER INFORMATION: coding sequence as in SEQ ID NO:1
US-09-988-626-3

Query Match 100.0%; Score 2958; DB 9; Length 2958;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2958; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCGCGGCGTAGTGACCGCGCTTTCTCAGTTTTGGTGACGCGGCGCATGTGGCGC 60
Db 1 CCGCGGCGTAGTGACCGCGCTTTCTCAGTTTTGGTGACGCGGCGCATGTGGCGC 60

Qy 61 TTGTGCTGCTGCTGCGGTCGCGCGCGGCGGACGACCATGTGCGAGGACGACCATATCG 120

Db	61	TTTTCTCGCTGCTCGGTGCTCGCGCCGGACGACCATGTGCGAGGAGCCCATATCGC	120
Qy	121	AGGCACCCGCGCGCGAGCGCGCAAGAACCCGCTGGCGCACCTGCGCGCAGCGGAG	180
Db	121	AGGCACCCGCGCGCGAGCGCGCGCAAGAACCCGCTGGCGCACCTGCGCGCAGCGGAG	180
Qy	181	AGAAAGCGGACCGTTCGGGGTGCTCCGGCGGGCCCAAAACACCGTGATCTGCGAGTGGTGG	240
Db	181	AGAAAGCGGACCGTTCGGGGTGCTCCGGCGGGCCCAAAACACCGTGATCTGCGAGTGGTGG	240
Qy	241	CAGCGGTAGCCGGGACTCGGGCGCCGCGCTCTACGTCTTTCCGAGTTCAACCGGTATC	300
Db	241	CAGCGGTAGCCGGGACTCGGGCGCCGCGCTCTACGTCTTTCCGAGTTCAACCGGTATC	300
Qy	301	TCCTTCAACTGTGGAAGGCGTTTCAGAGACTCATGCAAGGACACAAAGTTAAAGTTGCTC	360
Db	301	TCCTTCAACTGTGGAAGGCGTTTCAGAGACTCATGCAAGGACACAAAGTTAAAGTTGCTC	360
Qy	361	GCCTGGACAAATATTCCTGACACGAATGCACCTGGTCTTAATGTTGGGGCTTAAAGTGGAA	420
Db	361	GCCTGGACAAATATTCCTGACACGAATGCACCTGGTCTTAATGTTGGGGCTTAAAGTGGAA	420
Qy	421	TGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTTGACTTTCTGACACTCCACAAC	480
Db	421	TGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTTGACTTTCTGACACTCCACAAC	480
Qy	481	TGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGAATAGAACTGG	540
Db	481	TGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGAATAGAACTGG	540
Qy	541	CTGTGCGGCCCTACTCTGCCCCAGAAATACAGAGATGAAACCATGACAGTTTACCAGATCC	600
Db	541	CTGTGCGGCCCTACTCTGCCCCAGAAATACAGAGATGAAACCATGACAGTTTACCAGATCC	600
Qy	601	CCATACACAGTGAACAGAGAGGGGAAAGCAACCAACATGGCAGAGTCCAGAAAGGCCCTC	660
Db	601	CCATACACAGTGAACAGAGAGGGGAAAGCAACCAACATGGCAGAGTCCAGAAAGGCCCTC	660
Qy	661	TCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTGCAGTGAATAATGAGCCACAC	720
Db	661	TCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTGCAGTGAATAATGAGCCACAC	720
Qy	721	TTCCACATGTTGTAGCCAGAGAGAGGGTCAAGGACTCTTCCCTGCTGCTGCTTTCATCA	780
Db	721	TTCCACATGTTGTAGCCAGAGAGAGGGTCAAGGACTCTTCCCTGCTGCTGCTTTCATCA	780
Qy	781	TCCTGAAGCTTCACTTAAAGAGAGGAAACTCTTGGTGTCTCAAGCAAGGAGATGGGCC	840
Db	781	TCCTGAAGCTTCACTTAAAGAGAGGAAACTCTTGGTGTCTCAAGCAAGGAGATGGGCC	840
Qy	841	TCCAGTTGGGACAGCTGCCATCGCTCCCATCATTTGCTGCTGTCGAAGGACGGGAAAAAGCA	900
Db	841	TCCAGTTGGGACAGCTGCCATCGCTCCCATCATTTGCTGCTGTCGAAGGACGGGAAAAAGCA	900
Qy	901	TCACATCATGAAGACAGAGATTTTGGCTGAAGAGCTGTGACTCTCCTCCAGATCCTGGTG	960
Db	901	TCACATCATGAAGACAGAGATTTTGGCTGAAGAGCTGTGACTCTCCTCCAGATCCTGGTG	960
Qy	961	CTGCTTTTGTGGTGTAGATGTCAGATGAAAGCTTCATTCAACCCCATCTGTGAGAAATG	1020
Db	961	CTGCTTTTGTGGTGTAGATGTCAGATGAAAGCTTCATTCAACCCCATCTGTGAGAAATG	1020
Qy	1021	CCACCTTTCAGAGTACCAAGGAAAGGACAGATGCCCCGTGGCCCTTGTGTTTCAATGAG	1080
Db	1021	CCACCTTTCAGAGTACCAAGGAAAGGACAGATGCCCCGTGGCCCTTGTGTTTCAATGAG	1080
Qy	1081	CCCCAGCATCTGTGCTTGTGGACAGCAGAGTACACAGCTGGATGGAGAGTTTGGGGCTG	1140
Db	1081	CCCCAGCATCTGTGCTTGTGGACAGCAGAGTACACAGCTGGATGGAGAGTTTGGGGCTG	1140
Qy	1141	ACACCCAGCATTTGGTCTGGAATGAAACTGTGCTCTCAGTTTCAACAACCTTCGCAACCA	1200

Qy	2281	GAGTTGCTTTGACCA	CATGAAGGTC	TGCTTTGGAGA	CTTTCCAA	CAATGCCCAAGCTGA	2340
Db	2281	GAGTTGCTTTGACCA	CATGAAGGTC	TGCTTTGGAGA	CTTTCCAA	CAATGCCCAAGCTGA	2340
Qy	2341	TTCCCCACTGAAGAG	CCCTGTTTGTG	CGCATCGAGAGAT	CGAGAGCGAGG	GAGGAGGAGGAGA	2400
Db	2341	TTCCCCACTGAAGAG	CCCTGTTTGTG	CGCATCGAGAGAT	CGAGAGCGAGG	GAGGAGGAGGAGA	2400
Qy	2401	AGCGGAGCTGGCGCA	GGTGC	CGCGGCCCTCT	CTGTCCAGGGAGCT	GGCAGCGCGCTCG	2460
Db	2401	AGCGGAGCTGGCGCA	GGTGC	CGCGGCCCTCT	CTGTCCAGGGAGCT	GGCAGCGCGCTCG	2460
Qy	2461	AGGATGGGAGCCTC	AGCAGAAGCGGG	CCCCACACAGAGAG	GCACAGGCCACAGGCC	ACAGAGTCA	2520
Db	2461	AGGATGGGAGCCTC	AGCAGAAGCGGG	CCCCACACAGAGAG	GCACAGGCCACAGGCC	ACAGAGTCA	2520
Qy	2521	GAGCCAGTGAAGAT	CTGGGAGACCT	GAACTCAGAAGGCT	GTGTCTCTTC	TGCCCCCAGC	2580
Db	2521	GAGCCAGTGAAGAT	CTGGGAGACCT	GAACTCAGAAGGCT	GTGTCTCTTC	TGCCCCCAGC	2580
Qy	2581	CACGCACCCGTAT	CTGCCCTCTT	CTGTGTAGAGCT	TGAAGCAGAGAC	ACGTCGCCCAGGAGG	2640
Db	2581	CACGCACCCGTAT	CTGCCCTCTT	CTGTGTAGAGCT	TGAAGCAGAGAC	ACGTCGCCCAGGAGG	2640
Qy	2641	CAGCTCAGATAGGT	GGTATGAGCT	GTGCCGAGGCT	TGGGCTCCCA	ATAGCACTAGT	2700
Db	2641	CAGCTCAGATAGGT	GGTATGAGCT	GTGCCGAGGCT	TGGGCTCCCA	ATAGCACTAGT	2700
Qy	2701	CTATAGATGCCT	CTTAGGACT	GTGTCTGGCAC	CGCGGGCCAGAGG	GTGCCACACG	2760
Db	2701	CTATAGATGCCT	CTTAGGACT	GTGTCTGGCAC	CGCGGGCCAGAGG	GTGCCACACG	2760
Qy	2761	GAAGCAAGCAGAT	GAACAA	TTTCAATTTCAAGG	CAGTTTTTAAAGAA	GTCTTGGAAAACAG	2820
Db	2761	GAAGCAAGCAGAT	GAACAA	TTTCAATTTCAAGG	CAGTTTTTAAAGAA	GTCTTGGAAAACAG	2820
Qy	2821	ACGGCGCACCTTT	CTCTAATCCAGCA	AAAGTATCCCTGC	CACACCGAGAC	GAAGCAGA	2880
Db	2821	ACGGCGCACCTTT	CTCTAATCCAGCA	AAAGTATCCCTGC	CACACCGAGAC	GAAGCAGA	2880
Qy	2881	GTAAACAGGATCAG	TGGGTCTAAGT	GTCCGAGACTTAA	CGAAAAATAGTAT	TTTCAGCTC	2940
Db	2881	GTAAACAGGATCAG	TGGGTCTAAGT	GTCCGAGACTTAA	CGAAAAATAGTAT	TTTCAGCTC	2940
Qy	2941	TAAAGATTGAGTT	TGCAA	2958			
Db	2941	TAAAGATTGAGTT	TGCAA	2958			

RESULT 2

```

US-09-988-687-3
; Sequence 3, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE OF INVENTION: Gene and a Patalog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,687
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0

```

```

; SEQ ID NO 3
; LENGTH: 2958
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51)..(2531)
; OTHER INFORMATION: coding sequence as in SEQ ID NO:1
US-09-988-687-3

Query Match      100.0%; Score 2958; DB 9; Length 2958;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2958; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGGCGTAGGTGACCGGGCGCTTCTCAGTCTTTGTGGAGACGGCGCATGTGGGGCG 60
Db 1 CGCGGCGTAGGTGACCGGGCGCTTCTCAGTCTTTGTGGAGACGGCGCATGTGGGGCG 60

Qy 61 TTTGCTCGCTGCTGGGTCCGGGCGGAGCCACCATGTTCGAGGAGCGCACCATATCGC 120
Db 61 TTTGCTCGCTGCTGGGTCCGGGCGGAGCCACCATGTTCGAGGAGCGCACCATATCGC 120

Qy 121 AGGCACCGCCGCGCGAGCGCGGCCAAGGACCGCTCGGCACCTCGCGACGCGAG 180
Db 121 AGGCACCGCCGCGCGAGCGCGGCCAAGGACCGCTCGGCACCTCGCGACGCGAG 180

Qy 181 AGAAGCGCGGACCGTCGGGGTGCTCCGGCGGCCCAACACCGTGTACCTGCAGGTGTGG 240
Db 181 AGAAGCGCGGACCGTCGGGGTGCTCCGGCGGCCCAACACCGTGTACCTGCAGGTGTGG 240

Qy 241 CAGCGGTAGCCGGGACTCGGGCGCGCGCTCTACGTCTTCTCGAGTTCAACCGGTATC 300
Db 241 CAGCGGTAGCCGGGACTCGGGCGCGCGCTCTACGTCTTCTCGAGTTCAACCGGTATC 300

Qy 301 TCTTCAACTGTGGAGAGCGCTTCAGAGACTCATCGAGGACACAAGTTAAAGGTTGCTC 360
Db 301 TCTTCAACTGTGGAGAGCGCTTCAGAGACTCATCGAGGACACAAGTTAAAGGTTGCTC 360

Qy 361 GCCTGGACAAACATATTCCTGACAGCAATGCACTGGTCTAAATGTTGGGGGCTTAAAGTGAA 420
Db 361 GCCTGGACAAACATATTCCTGACAGCAATGCACTGGTCTAAATGTTGGGGGCTTAAAGTGAA 420

Qy 421 TGATTCTTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGACCTCCACAAC 480
Db 421 TGATTCTTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGACCTCCACAAC 480

Qy 481 TGGAAAATACTCGAGCAATCAAAATATTTCTGGTCCATTGAAAGGAATAGAACTGG 540
Db 481 TGGAAAATACTCGAGCAATCAAAATATTTCTGGTCCATTGAAAGGAATAGAACTGG 540

Qy 541 CTGTGCGGCCCACTCTGCCCCAGAAACAGAGGATGAAACCATGACAGTTTACCAAGATCC 600
Db 541 CTGTGCGGCCCACTCTGCCCCAGAAACAGAGGATGAAACCATGACAGTTTACCAAGATCC 600

Qy 601 CCATACACAGTGAACAGAGGAGGGGAAAGCAACCAATGCGCAGAGTCCGAAAGGCGCTC 660
Db 601 CCATACACAGTGAACAGAGGAGGGGAAAGCAACCAATGCGCAGAGTCCGAAAGGCGCTC 660

Qy 661 TCAGCAGGCTCAGTCCAGAGGCACTTCACAGACTCGGAGTCAATGAAATGAGCCACACC 720
Db 661 TCAGCAGGCTCAGTCCAGAGGCACTTCACAGACTCGGAGTCAATGAAATGAGCCACACC 720

Qy 721 TTCCACATGGTGTAGCCAGAGAGAGGGGTCAAGGACTCTTCCCTGGTCTGAGCTTTCA 780
Db 721 TTCCACATGGTGTAGCCAGAGAGAGGGGTCAAGGACTCTTCCCTGGTCTGAGCTTTCA 780

Qy 781 TCTGTAAGCTTCACTTAAAGAGAGAAACTTCTTGGTGCTCAAAGCAAGGAGATGGGCC 840
Db 781 TCTGTAAGCTTCACTTAAAGAGAGAAACTTCTTGGTGCTCAAAGCAAGGAGATGGGCC 840

Qy 841 TCCAGTTGGGACAGCTGCCATCGCTCCCATCATTTGCTGTCTCAAGGACGGGAAAGCA 900
Db 841 TCCAGTTGGGACAGCTGCCATCGCTCCCATCATTTGCTGTCTCAAGGACGGGAAAGCA 900

```

Qy 901 TCACATCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTCCAGATCCTGGT 960
Db 901 TCACATCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTCCAGATCCTGGT 960
Qy 961 CTGCTTTTGTGTGTGTAGAAATCTCCAGATGAAGAGCTTCATTCAACCCCATCTGTGAGAAATG 1020
Db 961 CTGCTTTTGTGTGTGTAGAAATCTCCAGATGAAGAGCTTCATTCAACCCCATCTGTGAGAAATG 1020
Qy 1021 CCACCTTTTCAGAGGTACCAAGAAAGGCAAGATGCCCGGTGGCTTGTGTGTTCACATGG 1080
Db 1021 CCACCTTTTCAGAGGTACCAAGAAAGGCAAGATGCCCGGTGGCTTGTGTGTTCACATGG 1080
Qy 1081 CCCAGCATCTGTCTGTGTGGACAGAGGTACACAGAGTGGATGAGAGGTTTGGGCGTG 1140
Db 1081 CCCAGCATCTGTCTGTGTGGACAGAGGTACACAGAGTGGATGAGAGGTTTGGGCGTG 1140
Qy 1141 ACACCCAGCACCTTGTCTGGAATGAGAACTGTGCTCAGTTTCACAACCTTTCGACGCCACA 1200
Db 1141 ACACCCAGCACCTTGTCTGGAATGAGAACTGTGCTCAGTTTCACAACCTTTCGACGCCACA 1200
Qy 1201 AGATTCAAACCCAGTCAACCTCATCCACCCGGACATCTTCCCTGTCTCACAGTTTCC 1260
Db 1201 AGATTCAAACCCAGTCAACCTCATCCACCCGGACATCTTCCCTGTCTCACAGTTTCC 1260
Qy 1261 GCTGTAAGAGAGAGGGCCCCACCTCAGTGTGCCCATGGTTCCAGGTGAATGCCCTCCTCA 1320
Db 1261 GCTGTAAGAGAGAGGGCCCCACCTCAGTGTGCCCATGGTTCCAGGTGAATGCCCTCCTCA 1320
Qy 1321 AGTACAGCTCCGTCCAGGAGGGAGTGGCAGAGGATGCCATTATTCTTCAATCCTG 1380
Db 1321 AGTACAGCTCCGTCCAGGAGGGAGTGGCAGAGGATGCCATTATTCTTCAATCCTG 1380
Qy 1381 AGGAATTCAATAGTTGAGGCGTGTGAGTTTCCCACTTCCAGCAGCGGTGCGAGGAGTACA 1440
Db 1381 AGGAATTCAATAGTTGAGGCGTGTGAGTTTCCCACTTCCAGCAGCGGTGCGAGGAGTACA 1440
Qy 1441 GGAGAGTCCGACAGACGGCCCCAGCCCGACAGAGAAAGAGTCACTACCCAGAAATCA 1500
Db 1441 GGAGAGTCCGACAGACGGCCCCAGCCCGACAGAGAAAGAGTCACTACCCAGAAATCA 1500
Qy 1501 TCTTCTTTGGAACAGGGTCTGCGATCCGATGAAGATTGGAATGTCAGTGCACACTTG 1560
Db 1501 TCTTCTTTGGAACAGGGTCTGCGATCCGATGAAGATTGGAATGTCAGTGCACACTTG 1560
Qy 1561 TCAACATAAGCCCCACACGCTCTCTGCTACTGGAATGTTGGTGGGCAATTTGGGCGAGC 1620
Db 1561 TCAACATAAGCCCCACACGCTCTCTGCTACTGGAATGTTGGTGGGCAATTTGGGCGAGC 1620
Qy 1621 TGTGCGGTCAATTACGAGACAGGTGGACAGGGTCTTGGGCAACCTTGGCTGTGTGTTG 1680
Db 1621 TGTGCGGTCAATTACGAGACAGGTGGACAGGGTCTTGGGCAACCTTGGCTGTGTGTTG 1680
Qy 1681 TGTCCCACTGCACGACAGATCACACACGGGCTTCCCAAGTATCTTGTGCAGAGAGAAC 1740
Db 1681 TGTCCCACTGCACGACAGATCACACACGGGCTTCCCAAGTATCTTGTGCAGAGAGAAC 1740
Qy 1741 GCGCTTGGCATCTTTTGGAAAGCCGCTTCACTTCTGTGTGTGTTGCCCTCCCAACCCAGC 1800
Db 1741 GCGCTTGGCATCTTTTGGAAAGCCGCTTCACTTCTGTGTGTGTTGCCCTCCCAACCCAGC 1800
Qy 1801 TCAAGAGCTGGCTCCAGCATACACACAGTGCAGAGGTCTTGCACCATCAGTACAGTA 1860
Db 1801 TCAAGAGCTGGCTCCAGCATACACACAGTGCAGAGGTCTTGCACCATCAGTACAGTA 1860
Qy 1861 TGATTCTGCCAAATGCTTTCAGGAAGGGGTGAGATCTCCAGTCTCTCAGTGGGAAAGAT 1920
Db 1861 TGATTCTGCCAAATGCTTTCAGGAAGGGGTGAGATCTCCAGTCTCTCAGTGGGAAAGAT 1920
Qy 1921 TGATCAGTTCTGTGTGGCAACATGTGATTTGGAAAGAGTTTCAGACCTGTCTGTGTGGCGC 1980
Db 1921 TGATCAGTTCTGTGTGGCAACATGTGATTTGGAAAGAGTTTCAGACCTGTCTGTGTGGCGC 1980

RESULT 3

US-09-988-626-223

; Sequence 223, Application US/09988626

; Publication No. US20030044959A1

Qy 1981 ACTCAAGCATGCGTTTGGCTGTGCGTGTGGTGACACCTCTGGCTGGAAAGTGTCTATT 2040
Db 1981 ACTCAAGCATGCGTTTGGCTGTGCGTGTGGTGACACCTCTGGCTGGAAAGTGTCTATT 2040
Qy 2041 CCGGGGACACCATCCCTTGCAGGCTCTGTGTCGGATGGGAAAGATGCCACCTCTCTGA 2100
Db 2041 CCGGGGACACCATCCCTTGCAGGCTCTGTGTCGGATGGGAAAGATGCCACCTCTCTGA 2100
Qy 2101 TACATGAAGCCACCTCGAAGATGGTTTGAAGAGGAGCAGTGGAAAGACACACAGCA 2160
Db 2101 TACATGAAGCCACCTCGAAGATGGTTTGAAGAGGAGCAGTGGAAAGACACACAGCA 2160
Qy 2161 CAACGTCCCAAGCCATCAGCGTGGGATCGGATGAACGGGAGTTCAATTATGCTGAAC 2220
Db 2161 CAACGTCCCAAGCCATCAGCGTGGGATCGGATGAACGGGAGTTCAATTATGCTGAAC 2220
Qy 2221 ACTTCAGCAGCGCTATGCCAAGGTCCCTCTTTCAGCCCCCACTTCAGCGAGAAAGTGG 2280
Db 2221 ACTTCAGCAGCGCTATGCCAAGGTCCCTCTTTCAGCCCCCACTTCAGCGAGAAAGTGG 2280
Qy 2281 GAGTTGCTTTGACCAATGAAGGTCTGCTTGGAGACTTTCACAACTTCCCAAGCTGA 2340
Db 2281 GAGTTGCTTTGACCAATGAAGGTCTGCTTGGAGACTTTCACAACTTCCCAAGCTGA 2340
Qy 2341 TTCCCCCAGTGAAGCCCTGTTTCTGGCGACATCGAGGAGATGGAGCGCAGGGAGA 2400
Db 2341 TTCCCCCAGTGAAGCCCTGTTTCTGGCGACATCGAGGAGATGGAGCGCAGGGAGA 2400
Qy 2401 AGCGGAGCTGCGGACAGTGGCGGCGCTCTCTGTCTCAGGAGCTGGCAGCGGCTGG 2460
Db 2401 AGCGGAGCTGCGGACAGTGGCGGCGCTCTCTGTCTCAGGAGCTGGCAGCGGCTGG 2460
Qy 2461 AGGATGGGAGCTCAGCAGAAAGCGGCCCTCCTGTCTCAGGAGCTGGCAGCGGCTGG 2520
Db 2461 AGGATGGGAGCTCAGCAGAAAGCGGCCCTCCTGTCTCAGGAGCTGGCAGCGGCTGG 2520
Qy 2521 GAGCCAGTGAAGATCTGGGAGACCTTGAACCTCAGAGGCTGTGTCTTCTGCCCCACG 2580
Db 2521 GAGCCAGTGAAGATCTGGGAGACCTTGAACCTCAGAGGCTGTGTCTTCTGCCCCACG 2580
Qy 2581 CACGCAACCGTATCTGCGCTCTTCTGCTGAGAGCTGAAAGACACCGTCCCCCAGGAGG 2640
Db 2581 CACGCAACCGTATCTGCGCTCTTCTGCTGAGAGCTGAAAGACACCGTCCCCCAGGAGG 2640
Qy 2641 CAGCTCAGGATAGGTGTATGGAGCTGTGCGAGGCTTGGGCTCCACATAAGCACTAGT 2700
Db 2641 CAGCTCAGGATAGGTGTATGGAGCTGTGCGAGGCTTGGGCTCCACATAAGCACTAGT 2700
Qy 2701 CTATAGATGCTCTTAGGACTGTGCTGCGACAGCGCGGCGCAGGAGGCTGCCACAG 2760
Db 2701 CTATAGATGCTCTTAGGACTGTGCTGCGACAGCGCGGCGCAGGAGGCTGCCACAG 2760
Qy 2761 GAAGCAAGCAGATGAATTAATTTCAAGGAGTCTTAAAGAGTCTTGAAGACAG 2820
Db 2761 GAAGCAAGCAGATGAATTAATTTCAAGGAGTCTTAAAGAGTCTTGAAGACAG 2820
Qy 2821 ACGCGGCACTTTCTCTAATCCAGCAAGTGTATCCCTGCACACAGAGACAAAGCAGA 2880
Db 2821 ACGCGGCACTTTCTCTAATCCAGCAAGTGTATCCCTGCACACAGAGACAAAGCAGA 2880
Qy 2881 GTAAAGGATCAGTGGGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGCTCAA 2940
Db 2881 GTAAAGGATCAGTGGGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGCTCAA 2940
Qy 2941 TAAAGATTGAGTTTGCAG 2958
Db 2941 TAAAGATTGAGTTTGCAG 2958


```
Db 1681 CAGAGAGAACGAGCCTTGCGATCTTTGGGAAAGCCCTTTACCCCTTTGCTGGTGTGCC 1740
Qy 1791 CCCAACCAAGCTCAAAGCCTGGCTCCAGCAGTACCAACACCAAGTGCAGAGGTCTCGAC 1850
Db 1741 CCCAACCAAGCTCAAAGCCTGGCTCCAGCAGTACCAACACCAAGTGCAGAGGTCTCGAC 1800
Qy 1851 CACATCAGTATGATTCTCGCCAAATGCCTTTCAGGAAGGGCTGAGATCTCCAGTCTCTCA 1910
Db 1801 CACATCAGTATGATTCTCGCCAAATGCCTTTCAGGAAGGGCTGAGATCTCCAGTCTCTCA 1860
Qy 1911 GTGGAAGATTGATCAGTTGCTGTGTGGCAACATGTGATTTTGGAAAGATTTTCAGACCTGT 1970
Db 1861 GTGGAAGATTGATCAGTTGCTGTGTGGCAACATGTGATTTTGGAAAGATTTTCAGACCTGT 1920
Qy 1971 CTGGTGGCACTCAAGCATGCGTTTGGCTGTGGCTGGTGCAACCTCTGGCTGGAAAA 2030
Db 1921 CTGGTGGCACTCAAGCATGCGTTTGGCTGTGGCTGGTGCAACCTCTGGCTGGAAAA 1980
Qy 2031 GTGGTCTATTCCGGGGACACCATGCCCTCGAGGCTCTGGTCCGGATGGGAAAGATGCC 2090
Db 1981 GTGGTCTATTCCGGGGACACCATGCCCTCGAGGCTCTGGTCCGGATGGGAAAGATGCC 2040
Qy 2091 ACCCTCTGATACATCAAGCCACCCCTGGAAAGATGTTTGGAAAGAGAAAGATGAAAAAG 2150
Db 2041 ACCCTCTGATACATCAAGCCACCCCTGGAAAGATGTTTGGAAAGAGAAAGATGAAAAAG 2100
Qy 2151 ACACACAGCACAACGTCCCAAGCCATCAGCTGGGGATGCGGATGAAACGGGAGTTCAAT 2210
Db 2101 ACACACAGCACAACGTCCCAAGCCATCAGCTGGGGATGCGGATGAAACGGGAGTTCAAT 2160
Qy 2211 ATGCTGAACCACTTCAGCAGGCGCTATGCCAAGTCCCTCTTCAGCCCACTTCAGC 2270
Db 2161 ATGCTGAACCACTTCAGCAGGCGCTATGCCAAGTCCCTCTTCAGCCCACTTCAGC 2220
Qy 2271 GAGAAAGTGGGAGTTGCCCTTTGACCACATGAAGTCTGCTTTGGAGACTTTTCCAACAATG 2330
Db 2221 GAGAAAGTGGGAGTTGCCCTTTGACCACATGAAGTCTGCTTTGGAGACTTTTCCAACAATG 2280
Qy 2331 CCCAAGCTGATTTCCCACTGAAAGCCTGTTGCTGGGACATCGAGAGATGAGGAG 2390
Db 2281 CCCAAGCTGATTTCCCACTGAAAGCCTGTTGCTGGGACATCGAGAGATGAGGAG 2340
Qy 2391 CGCAGGGAAGACGGGAGCTCGCGCAGTGGCGGGCGCCCTCTCTCCAGGAGCTGGCA 2450
Db 2341 CGCAGGGAAGACGGGAGCTCGCGCAGTGGCGGGCGCCCTCTCTCCAGGAGCTGGCA 2400
Qy 2451 GCGGCGCTGGAGATGGGAGCTTCAGCAGAAAGCGGGGCCACACAGAGAGCCACAGGCC 2510
Db 2401 GCGGCGCTGGAGATGGGAGCTTCAGCAGAAACGGGGCCACACAGAGAGCCACAGGCC 2460
Qy 2511 AAGAAAGTCAGAGCCAGTGAAGATCTGGGAGACCTGAACTCAGAAAGCTGTGTCTT 2570
Db 2461 AAGAAAGTCAGAGCCAGTGAAGATCTGGGAGACCTGAAATTCAGAAAGCTGTGTCTT 2520
Qy 2571 CTGCCCCCAGCAGCACCGTATCGCCCTCTGCTGTGTAGAACTGAAGAGCAGGTC 2630
Db 2521 CTGCCCCCAGCAGCACCGTATCGCCCTCTGCTGTGTAGAACTGAAGAGCAGGTC 2580
Qy 2631 CCCAGGAGGACAGCTCAGGATAGGTGATGAGAGCTGTGCCAGGCTTTGGGCTCCACAT 2690
Db 2581 CCCAGGAGGACAGCTCAGGATAGGTGATGAGAGCTGTGCCAAGCTTTGGGCTCCACAT 2640
Qy 2691 AAGCACTAGTCTATAGATGCTCTTAGGACTGTGCTGGGACACCGCGGGCCAGGAG 2750
Db 2641 AAGCACTAGTCTATAGATGCTCTTAGGACTGTGCTGGGACACCGCGGGCCAGGAG 2700
Qy 2751 CTGCCACACGGAAGCAAGCAGATGAATTAATTTCAAGSCAGTTTAAAGAGTC 2810
Db 2701 CTGCCACACGGAAGCAAGCAGATGAATTAATTTCAAGSCAGTTTAAAGAGTC 2760
Qy 2811 TTGGAACACAGCGGCGGACCTTTCTCTTAATCCAGCAAGTATTCCTGTCACACCCAGA 2870
Db 2761 TTGGAACACAGCGGCGGACCTTTCTCTTAATCCAGCAAGTATTCCTGTCACACCCAGA 2820
```

```
Qy 2871 GACAAGCAGAGTACAGGATCAGTGGGTCTAAGTGTCCGAGACTTAACGAAAAATAGTATT 2930
Db 2821 GACAAGCAGAGTACAGGATCAGTGGGTCTAAGTGTCCGAGACTTAACGAAAAATAGTATT 2880
Qy 2931 TCAGCTCAATAAAGATTGAGTTTGCAA 2958
Db 2881 TCAGCTCAATAAAGATTGAGTTTGCAA 2908
```

RESULT 4

```
US-09-988-687-223
; Sequence 223, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigan, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,687
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 223
; LENGTH: 2908
; TYPE: DNA
; ORGANISM: Pan troglodytes
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
US-09-988-687-223
```

```
Query Match 97.2%; Score 2874.4; DB 9; Length 2908;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 2887; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
```

```
Qy 51 ATGTGGGCGCTTTGCTCGCTGCTCGGTCCGCGCGGACGACCATGTGCGAGGACGC 110
Db 1 ATGTGGGCGCTTTGCTCGCTGCTCGGTCCGCGCGGACGACCATGTGCGAGGACGC 60
Qy 111 ACCATATCGCAGGACCCGCGCGGAGCGGCGGCGGCGCAAGNACCCGCTGCGGCACCTG 170
Db 61 ACCATATCGCAGGACCCGCGCGGAGCGGCGGCGGCGCAAGNACCCGCTGCGGCACCTG 120
Qy 171 CGCACGCGAGAGAACGCGGACCGTGGGTGCTCGGGCGGCCCAACACCGTGTAACCTG 230
Db 121 CGCACGCGAGAGAACGCGGACCGTGGGTGCTCGGGCGGCCCAACACCGTGTAACCTG 180
Qy 231 CAGGTGTGCGCAGCGGTAGCCGCGGACTCGGGCGCGCGCTCTACGCTCTTCTCCGAGTTC 290
Db 181 CAGGTGTGCGCAGCGGTAGCCGCGGACTCGGGCGCGCGCTCTACGCTCTTCTCCGAGTTC 240
Qy 291 AACCGGTATCTTCAACTGTGGAGAGGCGTTTCAGAGACTCATGCAAGGACCAAGTTA 350
Db 241 AACCGGTATCTTCAACTGTGGAGAGGCGTTTCAGAGACTCATGCAAGGACCAAGTTA 300
Qy 351 AAGGTGCTCGCTGGACAACATATTCCTGACACGAATGCACTGCTTAATGTTGGGGGC 410
Db 301 AAGGTGCTCGCTGGACAACATATTCCTGACACGAATGCACTGCTTAATGTTGGGGGC 360
Qy 411 TTAAGTGGAAATGATTTCTACTTTTAAAGGAAACCGGGCTTCCAAAGTGTGACTTTCTGGA 470
Db 361 TTAAGTGGAAATGATTTCTACTTTTAAAGGAAACCGGGCTTCCAAAGTGTGACTTTCTGGA 420
```

QY 471 CTTCCAACTGGAAAAATACTCGAAGCAATCAAAATATTTCTGGTCCAATTGAAAGCA 530
DB 421 CTTCCAACTGGAAAAATACTCGAAGCAATCAAAATATTTCTGGTCCAATTGAAAGCA 480
QY 531 ATAGAACTGGCTGTGGGGCCCACTCTGCCAGATACGAGATGAAACCAATGACAGTT 590
DB 481 ATAGAACTGGCTGTGGGGCCCACTCTGCCAGATACGAGATGAAACCAATGACAGTT 540
QY 591 TACCAGATCCCAATACACAGTGAACAGAGAGGGGAAAGCAACCAATGAGAGTCCA 650
DB 541 TACCAGATCCCAATACACAGTGAACAGAGAGGGGAAAGCAACCAATGAGAGTCCA 600
QY 651 GAAAGGCTCTCAGAGGCTCAGTCAGAGGAGATCTTCAGACTCCAGAGTGAATAAT 710
DB 601 GAAAGGCTCTCAGAGGCTCAGTCAGAGGAGATCTTCAGACTCCAGAGTGAATAAT 660
QY 711 GAGCCACACTTCCACATGGTGTAGCCAGAGAGGGGTGAGGACTCTTCCCTGGTC 770
DB 661 GAGCCACACTTCCACATGGTGTAGCCAGAGAGGGGTGAGGACTCTTCCCTGGTC 720
QY 771 GTAGCTTTTCATCTGTAACTTCACTTAAAGAGAGGAAACTTCTTGGTGTCTCAAGCAAG 830
DB 721 GTAGCTTTTCATCTGTAACTTCACTTAAAGAGAGGAAACTTCTTGGTGTCTCAAGCAAG 780
QY 831 GAGATGGGCTCCAGTTGGGACAGCTGCCATCGCTCCCATCATTCCTGCTGTCAAGGAC 890
DB 781 GAGATGGGCTCCAGTTGGGACAGCTGCCATCGCTCCCATCATTCCTGCTGTCAAGGAC 840
QY 891 GGGAAAGCATCACTCATGAAGAGAGAGATTTGGCTGAAGAGCTGTACTCTCTCA 950
DB 841 GGGAAAGCATCACTCATGAAGAGAGAGATTTGGCTGAAGAGCTGTACTCTCTCA 900
QY 951 GATCTGGTGTCTTTTGTGGGTGTAGATGTCCAGATGAAAGCTTCATTCAACCCATC 1010
DB 901 GATCTGGTGTCTTTTGTGGGTGTAGATGTCCAGATGAAAGCTTCATTCAACCCATC 960
QY 1011 TGTGAAATGCCACCTTTGAGAGGTACCAAGGAAAGGACAGATGCCCTTGGCTGGTG 1070
DB 961 TGTGAAATGCCACCTTTGAGAGGTACCAAGGAAAGGACAGATGCCCTTGGCTGGTG 1020
QY 1071 GTTCACTGGCCCCAGCATCTGTCTTGGACAGAGGATACAGAGCTGGATGGAGAG 1130
DB 1021 GTTCACTGGCCCCAGCATCTGTCTTGGACAGAGGATACAGAGCTGGATGGAGAG 1080
QY 1131 TTTGGGCTTGACACCCAGCACTTGTCTTGAATGAGAACTGTGCCCTCAGTTCAACACCTT 1190
DB 1081 TTTGGGCTTGACACCCAGCACTTGTCTTGAATGAGAACTGTGCCCTCAGTTCAACACCTT 1140
QY 1191 CGCAGCCAAAGATTCAAAACCCAGCTCAACCTCATCCACCGGACATCTTCCCTGCTC 1250
DB 1141 CGCAGCCAAAGATTCAAAACCCAGCTCAACCTCATCCACCGGACATCTTCCCTGCTC 1200
QY 1251 ACCAGTTTCGGCTGTAAGAGAGGGGCCACCTCAGTGTGCCATGGTTAGGGTGAA 1310
DB 1201 ACCAGTTTCGGCTGTAAGAGAGGGGCCACCTCAGTGTGCCATGGTTAGGGTGAA 1260
QY 1311 TGCTCTCAAGTACCAGCTCCGTCCAGAGGAGTGGCAGAGGATGCAATTATTACT 1370
DB 1261 TGCTCTCAAGTACCAGCTCCGTCCAGAGGAGTGGCAGAGGATGCAATTATTACT 1320
QY 1371 TGCAATCCTGAGGAATTCATGTAGGGGCTGCAGCTTCCCAACTTCCAGCAGAGCGTG 1430
DB 1321 TGCAATCCTGAGGAATTCATGTAGGGGCTGCAGCTTCCCAACTTCCAGCAGAGCGTG 1380
QY 1431 CAGGAGTACGAGGAGTGGCAGAGCGGCCAGCCCGCAGCAGAGAGAGAGTCAAGTAC 1490
DB 1381 CAGGAGTACGAGGAGTGGCAGAGCGGCCAGCCCGCAGCAGAGAGAGAGTCAAGTAC 1440
QY 1491 CCAGAAATCATCTTCTGGAAAGAGGCTGTGCCATCCCGATGAGATTGCAATGTCACT 1550
DB 1441 CCAGAAATCATCTTCTGGAAAGAGGCTGTGCCATCCCGATGAGATTGCAATGTCACT 1500

QY 1551 GCCACACTTGTCAACATTAAGCCCGACAGCTCTCTGTCTACTGGACTGTGGTGGAGGCACA 1610
DB 1501 GCCACACTTGTCAACATTAAGCCCGACAGCTCTCTGTCTACTGGACTGTGGTGGAGGCACG 1560
QY 1611 TTTGGGAGCTGTGCCGTCAATTACGAGACACAGGTGGACAGGCTCTGGGACACCTGGCT 1670
DB 1561 TTTGGGAGCTGTGCCGTCAATTACGAGACACAGGTGGACAGGCTCTGGGACACCTGGCT 1620
QY 1671 GCTGTGTTTGTGTCCACTCGACGAGATCACACACGGGCTTCCCAAGTATCTTGTGTG 1730
DB 1621 GCTGTGTTTGTGTCCACTCGACGAGATCACACACGGGCTTCCCAAGTATCTTGTGTG 1680
QY 1731 CAGAGAGAACGGGCTTGGCATCTTTGGGAAAGCGCTTACCCCTTGTGTGTGTGGC 1790
DB 1681 CAGAGAGAACGGGCTTGGCATCTTTGGGAAAGCGCTTACCCCTTGTGTGTGTGGC 1740
QY 1791 CCCAACACAGCTCAAAAGCTGGCTCCAGCAGTACCACACACAGTGCAGAGGTCTGTGCAC 1850
DB 1741 CCCAACACAGCTCAAAAGCTGGCTCCAGCAGTACCACACACAGTGCAGAGGTCTGTGCAC 1800
QY 1851 CACATCAGTATGATTTCTGCCAAATGCCTTCAGGAAGGGGCTGAGATCTCCAGTCTGTCA 1910
DB 1801 CACATCAGTATGATTTCTGCCAAATGCCTTCAGGAAGGGGCTGAGATCTCCAGTCTGTCA 1860
QY 1911 GTGGAAGATGATCAGTTGCTGTGGCAACATGTGATTTGGAAAGATTTTCAGACCTGT 1970
DB 1861 GTGGAAGATGATCAGTTGCTGTGGCAACATGTGATTTGGAAAGATTTTCAGACCTGT 1920
QY 1971 CTGGTGGGCACTGCAAGCATCGTTTGGCTGTGGCTGTGCACACCTCTGGCTGGGAA 2030
DB 1921 CTGGTGGGCACTGCAAGCATCGTTTGGCTGTGGCTGTGCACACCTCTGGCTGGGAA 1980
QY 2031 GTGGTCTATTCCGGGGACACCATGCTCGCAGGCTCTGGTCCGGATGGGAAAGATGCC 2090
DB 1981 GTGGTCTATTCCGGGGACACCATGCTCGCAGGCTCTGGTCCGGATGGGAAAGATGCC 2040
QY 2091 ACCCTCTGATACATGAAGCCACCTCGAAGATGTTTGGAAAGAGAGAGTGGGAAAG 2150
DB 2041 ACCCTCTGATACATGAAGCCACCTCGAAGATGTTTGGAAAGAGAGAGTGGGAAAG 2100
QY 2151 ACACACAGCACAACCTCCAGCCATCAGCGTGGGGATGCGGATGAACCGGAGTTCATT 2210
DB 2101 ACACACAGCACAACCTCCAGCCATCAGCGTGGGGATGCGGATGAACCGGAGTTCATT 2160
QY 2211 ATGCTGAAACCACTTCAGCCAGCGCTATGCCAAGTTCCTTTCAGCCCCCACTTCAGC 2270
DB 2161 ATGCTGAAACCACTTCAGCCAGCGCTATGCCAAGTTCCTTTCAGCCCCCACTTCAGC 2220
QY 2271 GAGAAAGTGGGAGTTGCTTTGACCAATGAAGTCTGCTTGGAGATCTTCCAACAATG 2330
DB 2221 GAGAAAGTGGGAGTTGCTTTGACCAATGAAGTCTGCTTGGAGATCTTCCAACAATG 2280
QY 2331 CCCAAGCTGATTTCCCACTGAAAGCCCTGTTGCTGGGACATCGAGAGATGAGGAG 2390
DB 2281 CCCAAGCTGATTTCCCACTGAAAGCCCTGTTGCTGGGACATCGAGAGATGAGGAG 2340
QY 2391 CGCAGGAGAAAGCGGAGCTGCGCAGGTGCGGGCGGCTCTCTTCCAGGAGCTGGCA 2450
DB 2341 CGCAGGAGAAAGCGGAGCTGCGCAGGTGCGGGCGGCTCTCTTCCAGGAGCTGGCA 2400
QY 2451 GCGGCTCTGGAGATGGGAGCTCAGCAGAAAGCGGGGCCACACAGAGAGCCACAGGCC 2510
DB 2401 GCGGCTCTGGAGATGGGAGCTCAGCAGAAAGCGGGGCCACACAGAGAGCCACAGGCC 2460
QY 2511 AAGAAAGTTCAGAGCCAGTGAAGATCTGGGACCCCTGAAATTCAGAAAGCTGTGTCTT 2570
DB 2461 AAGAAAGTTCAGAGCCAGTGAAGATCTGGGACCCCTGAAATTCAGAAAGCTGTGTCTT 2520
QY 2571 CTGCCCCCAGCAGCAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2630
DB 2521 CTGCCCCCAGCAGCAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2580
QY 2631 CCCAGGAGGAGCTCAGGATAGGTGATGAGGCTGTGCCGAGGCTTGGGCTCCACAT 2690

```
Db 2581 CCCAGGAGGAGCTCAGATAGTGGTATGAGCTGTGCCAAGGCTTGGGCTCCCAAT 2640
Qy 2691 AAGCACTAGTCTATAGATGCTCTTAGACTGGTGCCTGGCAAGCCGGGCCAGGAGG 2750
Db 2641 AAGCACTAGTCTATAGATGCTCTTAGACTGGTGCCTGGCAAGCCGGGCCAGGAGG 2700
Qy 2751 CTGCCACAGGAGCAAGCAGATGAATTAATTTTCAAGGCGAGTTTAAAGAGTC 2810
Db 2701 CTGCCACAGGAGCAAGCAGATGAATTAATTTTCAAGGCGAGTTTAAAGAGGC 2760
Qy 2811 TTGGAACACAGCCGGGACCTTCTCTTAATCCAGCAAGTGAATTCCTGCACACCAGA 2870
Db 2761 TTGGAACACAGCCGGGACCTTCTCTTAATCCAGCAAGTGAATTCCTGCACACCAGA 2820
Qy 2871 GACAGCAGAGTAACAGATCAGTGGGTCTAAGTGTCCGAGACTTAAACGAAATAGTATT 2930
Db 2821 GACAGCAGAGTAACAGATCAGTGGGTCTAAGTGTCCGAGACTTAAACGAAATAGTATT 2880
Qy 2931 TCAGCTGCAATAAAGATTGAGTTTGCAA 2958
Db 2881 TCAGCTGCAATAAAGATTGAGTTTGCAA 2908

RESULT 5
US-09-988-626-225
; Sequence 225, Application US/09988626
; Publication No. US20030044959A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,626
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 225
; LENGTH: 2892
; TYPE: DNA
; ORGANISM: Gorilla gorilla
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
US-09-988-626-225

Query Match 95.3%; Score 2819.6; DB 9; Length 2892;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 2863; Conservative 0; Mismatches 29; Indels 16; Gaps 1;

Qy 51 ATGTGGGCGCTTGTGCTGCTGCTGCGTCCGGCCGGAGCGGACCATGTTCGAGGGACGC 110
Db 1 ATGTGGGCGCTTGTGCTGCTGCTGCGTCCGGCCGGAGCGGACCATGTTCGAGGGACGC 60

Qy 111 ACCATATCGCAGGACCCCGCCCGCGAGCGGCGCGAAGGACCCGCTGCGGCACCTG 170
Db 61 ACCATATCGCAGGACCCCGCCCGCGAGCGGCGCGAAGGACCCGCTGCGGCACCTG 120

Qy 171 CGCAGCGAGAGAGCGGCGGACCTGCGGGTCTCGGGGGCCCAACACCGTGTACCTG 230
Db 121 CGCAGCGAGAGAGCGGCGGACCTGCGGGTCTCGGGGGCCCAACACCGTGTACCTG 180

Qy 231 CAGGTGGTGGCAGCGGGTAGCCGGGACTCGGGCGCGCGGCTCTACGCTCTTCTCCGAGTTC 290
```

```
Db 181 CAGTGTGGCAGCGGTAGCCGGACTCGGGCGCGGCTCTACGCTCTCTCCAGTTC 240
Qy 291 AACCGGTATCTCTTCAACTGTGGAGAGGCGTTTCCAGAGACTCATGACAGGACCAAGTTA 350
Db 241 AACCGGTATCTCTTCAACTGTGGAGAGGCGTTTCCAGAGACTCATGACAGGACCAAGTTA 300
Qy 351 AAGTTTCTCGCTGGGCAACATATTTCTCACAGGAATGCACTGGTCTAATGTGGGGGC 410
Db 301 AAGTTTCTCGCTGGGCAACATATTTCTCACAGGAATGCACTGGTCTAATGTGGGGGC 360
Qy 411 TTAAGTGAATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTTGA 470
Db 361 TTAAGTGAATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTTGA 420
Qy 471 CCTCCACAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGTGTCATTTGAAAGGA 530
Db 421 CCTCCACAGCTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGTGTCATTTGAAAGGA 480
Qy 531 ATAGAACTGGCTGTGCGGCCCACTCTGCCCCAGAAATACGAGGATGAACCATGACAGTT 590
Db 481 ATAGAACTGGCTGTGCGGCCCACTCTGCCCCAGAAATACGAGGATGAACCATGACAGTT 540
Qy 591 TACCAGATCCCCATACACAGTGAACAGAGAGGGAAGCAACCAACCATGGCAGAGTCCA 650
Db 541 TACCAGATCCCCATACACAGTGAACAGAGAGGGAAGCAACCAACCATGGCAGAGTCCA 600
Qy 651 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATCTTTCAGACTCCGAGTCAATGAAAT 710
Db 601 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATCTTTCAGACTCCGAGTCAATGAAAT 660
Qy 711 GAGCCACACTTCCACATGCTGTAGCCAGAGAGGAGGCTCAGGACTCTTCCCTGGTC 770
Db 661 GAGCCACACTTCCACATGCTGTAGCCAGAGAGGAGGCTCAGGACTCTTCCCTGGTC 720
Qy 771 GTAGCTTTTCACTGTAAAGCTTCACTTAAAGAGAGGAACTTCTTGGTGTCTCAAGCAAG 830
Db 721 GTAGCTTTTCACTGTAAAGCTTCACTTAAAGAGAGGAACTTCTTGGTGTCTCAAGCAAG 780
Qy 831 GAGATGGCTTCCAGTGGGACAGTGCATCGCTCCCATCATTTGCTGTGTCAAGAC 890
Db 781 GAGATGGCTTCCAGTGGGACAGTGCATCGCTCCCATCATTTGCTGTGTCAAGAC 840
Qy 891 GGGAAAGCATCACTCATGAAGNAGAGATTTTGGCTGAAGAGCTGTGTACTCTTCCA 950
Db 841 GGGAAAGCATCACTCATGAAGNAGAGATTTTGGCTGAAGAGCTGTGTACTCTTCCA 900
Qy 951 GATCCTGCTGCTGCTTTTGTGGTGAATGTCCAGATGAAAGCTTTCATTCAACCCATC 1010
Db 901 GATCCTGCTGCTGCTTTTGTGGTGAATGTCCAGATGAAAGCTTTCATTCAACCCATC 960
Qy 1011 TGTGAGAAATGCCACCTTTCAGAGGTCAAGGAAAGCAGATGCCCGCTGGCTTGGT 1070
Db 961 TGTGAGAAATGCCACCTTTCAGAGGTCAAGGAAAGCAGATGCCCGCTGGCTTGGT 1020
Qy 1071 GTTCACATGGCCCGCATCTGTGCTTGTGGACAGAGTACCAGCAGTGGATGGAGAGG 1130
Db 1021 GTTCACATGGCCCGCATCTGTGCTTGTGGACAGAGTACCAGCAGTGGATGGAGAGG 1080
Qy 1131 TTTGGGCTTGACACCCAGCACTTGGTCTCTGAATGAGAACTGTGCTCAGTTCAACACTT 1190
Db 1081 TTTGGGCTTGACACCCAGCACTTGGTCTCTGAATGAGAACTGTGCTCAGTTCAACACTT 1140
Qy 1191 CGCAGCACAAGATTCAAAACCCAGCTCAACCTCATCCACCGGACATCTTCCCGCTGCTC 1250
Db 1141 CGCAGCACAAGATTCAAAACCCAGCTCAACCTCATCCACCGGACATCTTCCCGCTGCTC 1200
Qy 1251 ACCAGTTTCCGCTGAAGAGAGGCGGCCCACTCAGTGTGCCATGGTTTCAAGGTTGAA 1310
Db 1201 ACCAGTTTCCGCTGAAGAGAGGCGGCCCACTCAGTGTGCCATGGTTTCAAGGTTGAA 1260
Qy 1311 TGCCTCTCAAGTACCAAGCTCCGCTCCAGGAGGAGTGGCAGAGGGATGCCATTATTACT 1370
```

Db 1261 TGCCTCCTCAAGTACCAGCTCCGTCCAGAGGGAAATGCGAGAGGATGCCATTATCACT 1320
QY 1371 TGAATCTGAGGAAATTCATAGTTGAGGCGCTGCACTTCCCAATTCACAGAGCGTG 1430
Db 1321 TGCAATCCTGAGGAAATTCATAGTTGAGGCGCTGCACTTCCCAATTCACAGAGTG 1380
QY 1431 CAGGAGTACAGGAGGAGTGGCAGGACGCGCCAGCCCGCAGCAGAGAAAGAGTCAGTAC 1490
Db 1381 CAGGAGTACAGGAGGAGTGTGAGGACGCTCCAGCCCGCAGCAGAGAAAGAGTCAGTAC 1440
QY 1491 CCAGAAATCATCTTCTTGGAAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCAGT 1550
Db 1441 CCAGAAATCATCTTCTTGGAAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCAGT 1500
QY 1551 GCCACACTTGTCAACATAAGCCCGCACAGTCTCTGCTACTGCACTGTGGTGAAGGACACA 1610
Db 1501 GCCACACTTGTCAACATAAGCCCGCACAGTCTCTGCTACTGCACTGTGGTGAAGGACAG 1560
QY 1611 TTTGGGACGCTGCGGTCATTACGAGACACAGGTGGAGGGTCTGGGCACTTGGCT 1670
Db 1561 TTTGGGACGCTGCGGTCATTACGAGACACAGGTGGAGGGTCTGGGCACTTGGCT 1620
QY 1671 GCTGTGTTTGTGTCACCTGCAAGCAGATCACACACGGGCTTGCCAAAGTATCTTGCTG 1730
Db 1621 GCTGTGTTTGTGTCACCTGCAAGCAGATCACACACGGGCTTGCCAAAGTATCTTGCTG 1680
QY 1731 CAGAGAGAACGGGCTTGGCATCTTTGGGAAAGCCGCTTCACTTGTGCTGTGGCT 1790
Db 1681 CAGAGAGAACGGGCTTGGCATCTTTGGGAAAGCCGCTTCACTTGTGCTGTGGCT 1740
QY 1791 CCCAACGCTCAAGCCTGGCTCAGAGTACCAACACAGTGCAGAGGTCTGTCAC 1850
Db 1741 CCCAACGCTCAAGCCTGGCTCAGAGTACCAACACAGTGCAGAGGTCTGTCAC 1800
QY 1851 CACATCAGTATGATTCCTGCCAAATCCCTTCCAGGAGGGCTGAGATCTCCAGTCTCGCA 1910
Db 1801 CACATCAGTATGATTCCTGCCAAATCCCTTCCAGGAGGGCTGAGATCTCCAGTCTCGCA 1860
QY 1911 GTGGAAGATTTGATCAGTTGCTGTTGCGAACAATGTGATTTGGAAGATTTTCCAGCTGT 1970
Db 1861 GTGGAAGATTTGATCAGTTGCTGTTGCGAACAATGTGATTTGGAAGATTTTCCAGCTGT 1920
QY 1971 CTGGTCCGGCACTGCAAGCATGGTTTGGCTGTGCTGTCGCTGGTGCACACCTCTGGCTGAAA 2030
Db 1921 CTGGTCCGGCACTGCAAGCATGGTTTGGCTGTGCTGTCGCTGGTGCACACCTCTGGCTGAAA 1980
QY 2031 GTGGTCTATTCCGGGACACCATGCTCGAGGCTCTGCTCGGATGGGAAAGTGC 2090
Db 1981 GTGGTCTATTCCGGGACACCATGCTCGAGGCTCTGCTCGGATGGGAAAGTGC 2040
QY 2091 ACCCTCCTGATACATGAAGCCACCTCGGAAGATGGTTTGGAGAGGAAGCAGTGGAAAAAG 2150
Db 2041 ACCCTCCTGATACATGAAGCCACCTCGGAAGATGGTTTGGAGAGGAAGCAGTGGAAAAAG 2100
QY 2151 ACACAGCACAACGTTCCCAAGCCATCAGCGTGGGATCGGATGAACCGGAGTTTCATT 2210
Db 2101 ACACAGCACAACGTTCCCAAGCCATCAGCGTGGGATCGGATGAACCGGAGTTTCATT 2160
QY 2211 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGTCCCTCTTCCAGCCCCAATTCAGC 2270
Db 2161 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGTCCCTCTTCCAGCCCCAATTCAGC 2220
QY 2271 GAGAAAGTGGGAGTTGCTTTGACCACTGAAGGTCTGCTTTGGAGACTTTTCCAAATG 2330
Db 2221 GAGAAAGTGGGAGTTGCTTTGACCACTGAAGGTCTGCTTTGGAGACTTTTCCAAATG 2280
QY 2331 CCCAAGCTGATTTCCCACTGAAGCCCTGTTTGTGTCGCAATCAGGAGATGGAGGAG 2390
Db 2281 CCCAAGCTGATTTCCCACTGAAGCCCTGTTTGTGTCGCAATCAGGAGATGGAGGAG 2340
QY 2391 CGCAGGAGAGCGGAGCTGCGGAGGTGCGGGCGGCGCTCTCTGTCAGGAGCTGGCA 2450
Db 2341 CGCAGGAGAGCGGAGCTGCGGAGGTGCGGGCGGCGCTCTCTGTCAGGAGCTGGCA 2400

QY 2451 GCGGCGCTTGGAGGATGGGAGCGCTTCCAGCAGAAAGCGGGCCCAACAGAGAGCCACAGGCC 2510
Db 2401 GCGGCGCTTGGAGGATGGGAGCGCTTCCAGCAGAAAGCGGGCCCAACAGAGAGCCACAGGCC 2460
QY 2511 AAGAGGTCAGAGGCCAGTGAAGATCTGGGAGACCTGAACTCAGAAAGCTGTGTGCTT 2570
Db 2461 AAGAGGTCAGAGGCCAGTGAAGATCTGGGAGACCTGAACTCAGAAAGCTGTGTGCTT 2520
QY 2571 CTGCCCCACGACGACCGCTATCTGCCCTCTCTTCTGCTGTAAGCTGAAGACGAGTGC 2630
Db 2521 CTGCCCCACGACGACCGCTATCTGCCCTCTCTTCTGCTGTAAGCTGAAGACGAGTGC 2580
QY 2631 CCCAGAGGAGGAGCTCAGGATAGGTGATGAGCTGTGCGGAGCTTGGGCTCCACAT 2690
Db 2581 CCCAGAGGAGGAGCTCAGGATAGGTGATGAGCTGTGCGGAGCTTGGGCTCCACAT 2640
QY 2691 AAGCACTAGTCTATAGATGCTCTTAGGACTGGTGCCTGGCAGCAGCGCGGCGCAGGAGG 2750
Db 2641 AAGCACTAGTCTATTA-----GGTGCCTGGCAGCAGCGCGGCGCAGGAGG 2684
QY 2751 CTGCCACACGGAAGCAAGCAGATGAATTAATTTCAATTTCAAGGCAAGTTTAAAGAAAGTC 2810
Db 2685 CTGCCACACGGAAGCAAGCAGATGAATTAATTTCAATTTCAAGGCAAGTTTAAAGAAAGTC 2744
QY 2811 TTGGAACACAGCGCGGACCTTCTCTAATCCAGCAAAAGTATTCCTGTCACACACAGA 2870
Db 2745 TTGGAACACAGCGCGGACCTTCTCTAATCCAGCAAAAGTATTCCTGTCACACACAGA 2804
QY 2871 GACAAGCAGATTAACAGATCAGTGGGTCTAAGTGTCCGAGACTTAAAGCAAAATAGTATT 2930
Db 2805 GACAAGCAGATTAACAGATCAGTGGGTCTAAGTGTCCGAGACTTAAAGCAAAATAGTATT 2864
QY 2931 TCAGCTGCAATAAAGATTGAGTTTGCAA 2958
Db 2865 TCAGCTGCAATAAAGATTGAGTTTGCAA 2892

RESULT 6

US-09-988-687-225
; Sequence 225, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,687
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 225
; LENGTH: 2892
; TYPE: DNA
; ORGANISM: Gorilla gorilla
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
US-09-988-687-225

Query Match 95.3%; Score 2819.6; DB 9; Length 2892;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 2863; Conservative 0; Mismatches 29; Indels 16; Gaps 1;


```
Db 2161 ATGCTGAACCACTTCAGCCAGGCTATGCAAGGTCCTCTTCAGCCCCAACTTCAAC 2220
Qy 2271 GAGAAAGTGGGAGTTCCTTTTACCACATGAAGGCTGCTTTGGAGACTTTTCCAAACAATG 2330
Db 2221 GAGAAAGTGGGAGTTCCTTTTACCACATGAAGGCTGCTTTGGAGACTTTTCCAAACAATG 2280
Qy 2331 CCCAAGCTGATTCCTCCCACTGAAGCCCTGTTTGTGGGAGCATCAGAGAGATGAGAGAG 2390
Db 2281 CCCAAGCTGATTCCTCCCACTGAAGCCCTGTTTGTGGGAGCATCAGAGAGATGAGAGAG 2340
Qy 2391 CCAGGGAGAACGGGAGCTCGGAGGTGCGGAGGTGCGGCGGCCCTCTGTCAGGAGCTGGCA 2450
Db 2341 CCAGGGAGAACGGGAGCTCGGAGGTGCGGAGGTGCGGCGGCCCTCTGTCAGGAGCTGGCA 2400
Qy 2451 GCGGCGCTGGAGATGGGAGGCTTCAGAGAAAGCGGGCCCAACAGAGAGAGCCAGAGGCC 2510
Db 2401 GCGGCGCTGGAGATGGGAGGCTTCAGAGAAAGCGGGCCCAACAGAGAGAGCCAGAGGCC 2460
Qy 2511 AAGAAAGTTCAGAGCCAGTGAAGATCTGGGAGACCTTGAACCTCAGAAAGCTGTGTCTT 2570
Db 2461 AAGAAAGTTCAGAGCCAGTGAAGATCTGGGAGACCTTGAACCTCAGAAAGCTGTGTCTT 2520
Qy 2571 CTGCCCCCAGCAGCCAGCTATCTGCCCTCTCTGCTGTAGAGCTGAAGCTGAAGCAGCGTC 2630
Db 2521 CTGCCCCCAGCAGCCAGCTATCTGCCCTCTCTGCTGTAGAGCTGAAGCTGAAGCAGCGTC 2580
Qy 2631 CCCCAGGAGGAGCTCAGATAGGTGGTATGAGAGCTGTCAGGAGCTTGGGCTCCACAT 2690
Db 2581 CCCCAGGAGGAGCTCAGATAGGTGGTATGAGAGCTGTCAGGAGCTTGGGCTCCACAT 2640
Qy 2691 AAGCACTAGCTATAGATGCTCTTAGGACTGGTGGCAGACCGCGGCCAGGAGG 2750
Db 2641 AAGCACTAGCTATAGATGCTCTTAGGACTGGTGGCAGACCGCGGCCAGGAGG 2684
Qy 2751 CTGCCCACAGGAGCAAGCAGATGACTAATTTCAATTTCAAGGAGTCTTTAAAGAGTC 2810
Db 2685 CTGCCCACAGGAGCAAGCAGATGACTAATTTCAATTTCAAGGAGTCTTTAAAGAGTC 2744
Qy 2811 TTGAAAACAGACGGCGGACCTTCTCTTAATCCAGCAAGTATTCCTCTGACACACCAGA 2870
Db 2745 TTGAAAACAGACGGCGGACCTTCTCTTAATCCAGCAAGTATTCCTCTGACACACCAGA 2804
Qy 2871 GACAGCAGAGTAACAGATCAGTGGTCTAAGTGTCCGAGACTTAACGAAAATAGTATT 2930
Db 2805 GACAGCAGAGTAACAGATCAGTGGTCTAAGTGTCCGAGACTTAACGAAAATAGTATT 2864
Qy 2931 TCAGTGCATTAAGTATGAGTTTGCAT 2958
Db 2865 TCAGTGCATTAAGTATGAGTTTGCAT 2892
```

RESULT 7

```
US-09-988-626-1
; Sequence 1, Application US/09988626
; Publication No. US2003004959A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; TITLE OF INVENTION: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,626
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
```

```
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2481
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
US-09-988-626-1
```

```
Query Match 83.9%; Score 2481; DB 9; Length 2481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 51 ATGTGGGCGCTTTGCTCCCTGCTGCGGTCCGCGCGCGAGCCGACCCGCTGCGGACCGC 110
Db 1 ATGTGGGCGCTTTGCTCCCTGCTGCGGTCCGCGCGCGAGCCGACCCGCTGCGGACCGC 60
Qy 111 ACCATATCCAGAGCAACCCGCGCGCGAGCGCGCCGCAAGGACCCGCTGCGGACCGTC 170
Db 61 ACCATATCCAGAGCAACCCGCGCGCGAGCGCGCCGCAAGGACCCGCTGCGGACCGTC 120
Qy 171 GCGAGCGAGAGACGCGGAGCCGTCGCGGTGCTCCGCGCGCGCCCAAAACACCGTGTACCTG 230
Db 121 GCGAGCGAGAGAGCGCGGAGCCGTCGCGGTGCTCCGCGCGCGCCCAAAACACCGTGTACCTG 180
Qy 231 CAGGTGGTGGCAGCGGCTAGCGGAGCTCGGCGCGCGCTCTACGCTCTTCTCCGAGTTC 290
Db 181 CAGGTGGTGGCAGCGGCTAGCGGAGCTCGGCGCGCGCTCTACGCTCTTCTCCGAGTTC 240
Qy 291 AACCGGTATCTTCAACTGTGGAGAGCGGCTTCAGAGACTCATCGAGAGACACAAGTTA 350
Db 241 AACCGGTATCTTCAACTGTGGAGAGCGGCTTCAGAGACTCATCGAGAGACACAAGTTA 300
Qy 351 AAGTTGCTCCGCTGGACAAACATATTCCTGACACAAATGCACTGCTCTAATGTTGGGGC 410
Db 301 AAGTTGCTCCGCTGGACAAACATATTCCTGACACAAATGCACTGCTCTAATGTTGGGGC 360
Qy 411 TTAAGTGAATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTACTTCTTCTGGA 470
Db 361 TTAAGTGAATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTACTTCTTCTGGA 420
Qy 471 CTTCCAACTCGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTTGAAAGGA 530
Db 421 CTTCCAACTCGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTTGAAAGGA 480
Qy 531 ATAGAACTGGCTGTGCGGCGCCCACTCTGCGCGCGAGGATGAAACCATGACAGTT 590
Db 481 ATAGAACTGGCTGTGCGGCGCCCACTCTGCGCGCGAGGATGAAACCATGACAGTT 540
Qy 591 TACCAGATCCCATACACAGTGAACAGAGAGGAGGAAAGCAACCACTGGCAGAGTCCA 650
Db 541 TACCAGATCCCATACACAGTGAACAGAGAGGAGGAAAGCAACCACTGGCAGAGTCCA 600
Qy 651 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCCGAATGAAAT 710
Db 601 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCCGAATGAAAT 660
Qy 711 GAGCCACACCTTCCACATGTTAGCCAGAGAGAGGGGTGAGGAGTCTTCCCTGTGTC 770
Db 661 GAGCCACACCTTCCACATGTTAGCCAGAGAGAGGGGTGAGGAGTCTTCCCTGTGTC 720
Qy 771 GTAGCTTTTCACTGTAAAGCTTCACTTAAAGAGAGAACTTCTTGGTGTCTCAAGCAAG 830
Db 721 GTAGCTTTTCACTGTAAAGCTTCACTTAAAGAGAGAACTTCTTGGTGTCTCAAGCAAG 780
Qy 831 GAGATGGGCTCCCACTGGGAGCAGCTGCCATCGCTCCCATCATTTGCTGCTCAAGGAC 890
Db 781 GAGATGGGCTCCCACTGGGAGCAGCTGCCATCGCTCCCATCATTTGCTGCTCAAGGAC 840
Qy 891 GGGAAAGCATCATCTGAGAGAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTCTCCA 950
Db 841 GGGAAAGCATCATCTGAGAGAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTCTCCA 900
```

Db 841 GGGAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTCTCCA 900
Qy 951 GATCTGTGTCTGCTTTTGTGTGTGTAGAAATGTCAGATGAAGAGCTTCAATCAACCCATC 1010
Db 901 GATCTGTGTCTGCTTTTGTGTGTGTAGAAATGTCAGATGAAGAGCTTCAATCAACCCATC 960
Qy 1011 TGTGAGAAATGCCATTTTCAAGAGTACCAAGGAAGGAGATGCCCGTGGCTTGGTG 1070
Db 961 TGTGAGAAATGCCATTTTCAAGAGTACCAAGGAAGGAGATGCCCGTGGCTTGGTG 1020
Qy 1071 GTTTCATGTGGCCCGCCAGCATCTGTGTGTGTGGACAGAGGTACCAAGAGGAGG 1130
Db 1021 GTTTCATGTGGCCCGCCAGCATCTGTGTGTGTGGACAGAGGTACCAAGAGG 1080
Qy 1131 TTTGGGCTGTACACCCAGCATTTGTGTCTGTGAATGAAATCTGTGCTCTAGTTTCACAACTT 1190
Db 1081 TTTGGGCTGTACACCCAGCATTTGTGTCTGTGAATGAAATCTGTGCTCTAGTTTCACAACTT 1140
Qy 1191 CGCAGCCACAAGATTCAAACCCAGCTCAACCTCACTCCACCCGGACATCTTCCCTCTGTC 1250
Db 1141 CGCAGCCACAAGATTCAAACCCAGCTCAACCTCACTCCACCCGGACATCTTCCCTCTGTC 1200
Qy 1251 ACCAGTTTCCGCTGTAGAAGAGGAGGCCCCACCTCTAGTGTGCCATGGTTTCAGGGTCAA 1310
Db 1201 ACCAGTTTCCGCTGTAGAAGAGGAGGCCCCACCTCTAGTGTGCCATGGTTTCAGGGTCAA 1260
Qy 1311 TGCCTCTCAAGTACAGCTCCGTCCAGAGGAGGAGTGGCAGAGGATGCCATATTACT 1370
Db 1261 TGCCTCTCAAGTACAGCTCCGTCCAGAGGAGGAGTGGCAGAGGATGCCATATTACT 1320
Qy 1371 TGCATCTCTAGGAAATTCATAGTTGAGGCGTGCAGCTTCCCACTTCCACGACAGCGTG 1430
Db 1321 TGCATCTCTAGGAAATTCATAGTTGAGGCGTGCAGCTTCCCACTTCCACGACAGCGTG 1380
Qy 1431 CAGGAGTACAGAGAGGTGCGCAGGACGGCCCGAGCCCGCAGCAGAGAGAAAGAGTCAGTAC 1490
Db 1381 CAGGAGTACAGAGAGGTGCGCAGGACGGCCCGAGCCCGCAGCAGAGAGAAAGAGTCAGTAC 1440
Qy 1491 CCAGAAATCATCTTCTTGGAAACAGGGTCTGCCATCCCGATCCCGATGGAATTCGAAATGTCAGT 1550
Db 1441 CCAGAAATCATCTTCTTGGAAACAGGGTCTGCCATCCCGATGGAATTCGAAATGTCAGT 1500
Qy 1551 GCCACACTTGTCAACATGAAGCCCGACACGCTCTGCTACTGGACTGTGGTGGGACACA 1610
Db 1501 GCCACACTTGTCAACATGAAGCCCGACACGCTCTGCTACTGGACTGTGGTGGGACACA 1560
Qy 1611 TTTGGGACAGCTGTGCGCTCATTTACGAGACAGGATGGACAGGTCCTGGGACCCCTGGCT 1670
Db 1561 TTTGGGACAGCTGTGCGCTCATTTACGAGACAGGATGGACAGGTCCTGGGACCCCTGGCT 1620
Qy 1671 GCTGTGTTTGTGTCACCTGCAACAGATCCACACAGGGTTCGCAAGTATCTTGTCTG 1730
Db 1621 GCTGTGTTTGTGTCACCTGCAACAGATCCACACAGGGTTCGCAAGTATCTTGTCTG 1680
Qy 1731 CAGAGAGAACGGCTTGGCATCTTTGGAAAGCGCTTACCCCTTGCCTGTGGTGTGCC 1790
Db 1681 CAGAGAGAACGGCTTGGCATCTTTGGAAAGCGCTTACCCCTTGCCTGTGGTGTGCC 1740
Qy 1791 CCCAACAGCTCAAAAGCTGTGCTCCAGCAGTACCAACACAGTGCAGAGGATCCTGCAC 1850
Db 1741 CCCAACAGCTCAAAAGCTGTGCTCCAGCAGTACCAACACAGTGCAGAGGATCCTGCAC 1800
Qy 1851 CACATCAGTATGATTCCTGCGCAATATGCTTCAGGAAGGGGTGAGATCTCAGTCTGTCA 1910
Db 1801 CACATCAGTATGATTCCTGCGCAATATGCTTCAGGAAGGGGTGAGATCTCAGTCTGTCA 1860
Qy 1911 GTGGAAGATTCAGTTCGCTGTGGCAATCATGTGATTTGGAGAGTTTCAGACCTGT 1970
Db 1861 GTGGAAGATTCAGTTCGCTGTGGCAATCATGTGATTTGGAGAGTTTCAGACCTGT 1920
Qy 1971 CTGCTGCGGCACTGCAAGCATGCTGTTGGCTGTGGCTGTGGTGCACACCTCTGGCTGGAAA 2030
Db 1921 CTGCTGCGGCACTGCAAGCATGCTGTTGGCTGTGGCTGTGGTGCACACCTCTGGCTGGAAA 1980

Qy 2031 GTGGTCTATTCCGGGACACCATGCGCTCGAGGCTCTGCTCGGATGGGAAGATGCC 2090
Db 1981 GTGGTCTATTCCGGGACACCATGCGCTCGAGGCTCTGCTCGGATGGGAAGATGCC 2040
Qy 2091 ACCCTCTGATATCATGAAGCACCTCTGGAAGATGGTTTGAAGAGGAGCAGTGGAAAAG 2150
Db 2041 ACCCTCTGATATCATGAGCCACCTCTGGAAGATGGTTTGAAGAGGAGCAGTGGAAAAG 2100
Qy 2151 ACACACAGCAACAGTCCCAAGCCATCAGCGTGGGATGCGGATGAACGCGGAGTTCAAT 2210
Db 2101 ACACACAGCAACAGTCCCAAGCCATCAGCGTGGGATGCGGATGAACGCGGAGTTCAAT 2160
Qy 2211 ATGCTGAACACTTCAGCAGCGCTATGCCAAGTCTCCCTCTTCAGCCCCCACTTCAGC 2270
Db 2161 ATGCTGAACACTTCAGCAGCGCTATGCCAAGTCTCCCTCTTCAGCCCCCACTTCAGC 2220
Qy 2271 GAGAAAGTGGGAGTTGCTTTTGACCAATGAAGTCTGCTTTGGAGACTTTCCAAACAATG 2330
Db 2221 GAGAAAGTGGGAGTTGCTTTTGACCAATGAAGTCTGCTTTGGAGACTTTCCAAACAATG 2280
Qy 2331 CCCAAGCTGATTTCCCACTGAAAGCCTGTTGCTGGCGACATCGAGAGATGGAGAG 2390
Db 2281 CCCAAGCTGATTTCCCACTGAAAGCCTGTTGCTGGCGACATCGAGAGATGGAGAG 2340
Qy 2391 CGCAGGAGAGGAGGAGTGGCGCAGTGGCGGCGGCGCTCTGCTCCAGGAGCTGGCA 2450
Db 2341 CGCAGGAGAGGAGGAGTGGCGCAGTGGCGGCGGCGCTCTGCTCCAGGAGCTGGCA 2400
Qy 2451 GCGGCGCTGAGGATGGGAGCCTCAGCAAGAGCGGCGCCACACAGAGGAGCCACAGGCC 2510
Db 2401 GCGGCGCTGAGGATGGGAGCCTCAGCAAGAGCGGCGCCACACAGAGGAGCCACAGGCC 2460
Qy 2511 AAGAAGTCCAGAGCCCAAGTGA 2531
Db 2461 AAGAAGTCCAGAGCCCAAGTGA 2481

RESULT 8
US-09-988-687-1
; Sequence 1, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigan, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988, 687
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2481
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
US-09-988-687-1

Query Match 83.9%; Score 2481; DB 9; Length 2481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 648 CCAGAAAGCCCTCTCAGCAGGCTCAGTCAGAGCGATCTTCAGATCCGAGTCGAATGAA 707
Db 574 CCCAGAAATCTCCCAACAGGCTCAGTCCCAACAGTATCGGATCTGGATCAGCTGAA 633
Qy 708 AATGAGCCACACTTCCACATCGGTGTAGCCAGAGAGAGGGGTCCAGGACTCTTCCCTG 767
Db 634 AATGGGC-----AGTCCACAGGAAGCATGGGGCAGGGAC-CCTCCTTA 678
Qy 768 GTCTGAGCTTTTCACTGTAAGCTTCACTTAAAGAGAGAACTTCTTGGTGTCAAAGCA 827
Db 679 GTGGTAGCTTTTGTCTGCAAGCTTCACTTGAGGAAGGAATCTTCTGGTCTTAAAGCA 738
Qy 828 AAGAGATGGGCTCCAGTTGGGACAGCTGCCATCGCTCCCATATCTGCTGCTCAAG 887
Db 739 AAGGAGCTGGGCTTCTGTTGGGACGCGCCCATTTGACCCCATATCTGCTGCTCAAG 798
Qy 888 GACGGGAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCT 947
Db 799 GACGGGAAGAGTATCACTTACGAGGAAGAGAGATTTGCTGAAGAGCTTTGTACACCC 858
Qy 948 CCAGATCTGTGTGCTGCTTTTGTGTGTGTAGAAATGTCAGATGAAAGTTCATTCAACCC 1007
Db 859 CCAGATCTGTGTGCTGCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 918
Qy 1008 ATCTGTGAAATGCCACCTTTTCAAGAGGTACCAAGAAAGGAGAGTCCCGGTGSCCTTG 1067
Db 919 ATCTGTGAAGACGACACCTTTTAAAGGTACCAAGGAGAGGTGATGCACTGTGGCGCTG 978
Qy 1068 GTGGTTTCAATGGGCCCCAGCATCTGTCTGTGGACAGCAGGTACACAGTGGATGGAG 1127
Db 979 GTGGTCCATAGCCCCAGATCTGTACTATCTGACAGCAGATACACAGTGGATGGAG 1038
Qy 1128 AGTTTGGGCTGTGACACCCAGCACTTGTCTGTAATGAGAACTGTGCTCTCACTTCAACAC 1187
Db 1039 AGTTTGGGCTGTGACACAGCAGCCTGTATCTGTAATGAGAAATTTGCCCTCGTCCACAAC 1098
Qy 1188 CTTCCAGGCCAAGAAATTCAAACCCAGCTCAACTCATCCACCGGACATCTTCCCCCTG 1247
Db 1099 CTGCGCAGCCCAAGAAATTCAGACCCAGCTCAGCCTCATCCACCCCTGACATCTTCCCCCAG 1158
Qy 1248 CTCACAGTTTCCGCTGTAAGAAAGAGGGGCCACCCCTCAGTGTGCCATGTCTCAGGCT 1307
Db 1159 CTTACACAGCTTCTAGTAAGAGAGAGAGGTCCACCTCAGGCTGCCAACAGTTTCGGGGT 1218
Qy 1308 GAATGCCCTCTCAAGTACAGCTCCGTCCAGAGGGAGTGGCAGAGGATGCCAATATT 1367
Db 1219 GAATGCCCTCTCAAGTATTTCAGTCCGCCCCCAAGAGAGAGTGGCAGAGGATACCAACTC 1278
Qy 1368 ACTTGCATCTGAGGAATTCATAGTTGAGGCGCTGCGAGCTTCCCACTTCCAGCAGAC 1427
Db 1279 GACTCAATACTGATGAATTCATAGCTGAGGCTTTGGAGCTCCCCAGTTTTCAGGAGAT 1338
Qy 1428 GTGAGGAGTACAGGAGGTGCGCAGGAGCGGCCCCAGCCCGCAGCAGAGAGAAAGAGTCAG 1487
Db 1339 GTGAGGAGTATCGAGAGAACTGTGAGAGAAACCCAGCCCGCAGCAGAGAAAGAGCCAG 1398
Qy 1488 TACCCAGAAATCATCTTCTTGGAAACAGGGTCTGCCATCCCGATGAAGATTGCAAAATGTC 1547
Db 1399 TATCTGAAATTTGCTTCTTGGGTACGGGTCTGCCATCCCAATCGAATCCGAAATGTC 1458
Qy 1548 AGTGCCACACTTGTCAACATAGCCCGCAGCAGCTCTCTGCTACTTGGATGTGGTGAAGGC 1607
Db 1459 AGTTCCACACTCGTCAACTAAGCCCTTACAAAGTCAGTGTCTGCTGATTTGGGAGAGGC 1518
Qy 1608 ACATTTGGGACAGCTGTGCGCTCATTCAGGAGACCAAGGTGGACAGGCTCTGGGCAACCTG 1667
Db 1519 ACTTTTGGGACAGTTGCGGCTCATTCGGCAGCAAAATAGACCCAGTCTTATGACGCTC 1578
Qy 1668 GCTGTGTGTTGTGCCACCTGTCCAGCAGATACCAACAGGGCTTCCCAAGTATCTTG 1727
Db 1579 ACGGCTGTGTTGTGCCACCTGTCCAGCAGCCGACCCACCAACAGGGCTTGTGTAATCTTG 1638

Qy 1728 CTGCAGAGAGAACCGCGCTTGGCATCTTTGGGAAGCGCTTACCCCTTTGCTGCTGTT 1787
Db 1639 CTGCAGAGAGAGCATGCGTTGGCATCTCTGGGAAAACCCCTTCCAGCCCTTCTTGTGTG 1698
Qy 1788 GCCCCCAACAGCTCAAAAGCCTGGCTCCAGCAGTACCACACCAAGTGCAGGAGGCTCTG 1847
Db 1699 GCTCTTACCAGCTCAGGGCTGGCTGCAGCAGTATCACAACCACTGCCAGGAGATTCTG 1758
Qy 1848 CACCATATAGTATGATTCCTGCAAAATGCTTCCAGAAAGGGCTGAGATCTCCAGTCT 1907
Db 1759 CACCACGCTCAGTATGATTCCTGCAAAATGCTTCCAGAAAGGGCAGAGGTCTCCAATACT 1818
Qy 1908 GCAGTGAAGATGATCAGTTCGCTGTGCAACATGTGATTTGGAAAGAGTTCAGAGC 1967
Db 1819 ACATTGGAAGGCTGATAAGCTTGTGTGAAACATGTGACTTTAGAAAGATTTTCAGACC 1878
Qy 1968 TGTCTGTGGGGCACTGCAAGCATGCTTTTGGCTGTGGCTGTGGTGCACACCTCTGGCTGG 2027
Db 1879 TGCCTGTGCGGCACTGCAAGCATGCTTTTGGCTGTGCTGTGCTGATTCATCTGGCTGG 1938
Qy 2028 AAAGTGTCTTATTCGGGGACACCATGCCCTGCAGGCTCTGTGCTCGGATGGGGAAGAT 2087
Db 1939 AAAGTCTCTACTCGGGGATACCATGCCCTGTGAGGCTCTGTGCTCCAGATGGGGAAGAT 1998
Qy 2088 GCCACCTCTGATACATGAAGCCACCTCGAAGATGTTTGGAAAGAGAGAGTGGAA 2147
Db 1999 GCCACCTCTGATACATGAAGCCACCTCTGAGGATCNCITTTGGAAGAGAGAGTGGAG 2058
Qy 2148 AAGACACACAGCACAACGCTCCCAAGCCATCAGCGTGGGGATGCGGATCAACGCGAGTTC 2207
Db 2059 AGGACACACAGCACCACCTCCAGGCTATTAAATGTGGGGATGCGGATGAATGCGAGTTC 2118
Qy 2208 ATTTATGCTGAACCACTTTCAGCAGCGCTATGCAAGGTCCCTCTTTCAGGCCCACTTC 2267
Db 2119 ATCATGTGTAACCACTTTCAGTCAAGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 2178
Qy 2268 AGCGAGAAAGTGGGAGTGGCTTTGACCATGAAGCTGCTTTGGAGACTTTTCCAAACA 2327
Db 2179 AACGAGAAAGTGGGATGCGCTTTGACCATGAAGTCTGNTTTGGAGACTTCCCGACA 2238
Qy 2328 ATGCCCAAGCTGATTTCCCACTGAAAGCCTGTTTCTGCGGACATCGAGGAGATGGAG 2387
Db 2239 GTGCCCAAGCTGATTTCCCACTGAAAGCCTGTTTTCAGGTGACATGAAGAGATGGTG 2298
Qy 2388 GAGCGCAGGAGAAAGCGGAGCTGCGGACAGGTGCGGGCGGCTCTCTGTCAGGGAGCTG 2447
Db 2299 GAACGCGAGGAGAAAGAGGAGCTACGGCTGTGTCGAGCAGGCTCTCTGACC---CAGCAG 2355
Qy 2448 GCAGCGCGCTGGAGATGGGAGCTTACGAGAAAGCGGGCCCAACAGAGAGGAGCCACA 2506
Db 2356 GCAGACAGCCAGAGGACAGAGAAACCCCAACAGAAAGCGGGCCCAACAGATGAACACCA 2414

RESULT 11

US-09-833-381-2039

; Sequence 2039, Application US/09833381

; Patent No. US20020132090A1

; GENERAL INFORMATION:

; APPLICANT: Robison, Keith E.

; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs

; FILE REFERENCE: 5800-119

; CURRENT APPLICATION NUMBER: US/09/833,381

; CURRENT FILING DATE: 2001-04-11

; PRIOR APPLICATION NUMBER: 09/516,448

; PRIOR FILING DATE: 2000-02-29

; NUMBER OF SEQ ID NOS: 2050

; SOFTWARE: Fast-Seq for Windows Version 3.0

; SEQ ID NO 2039

; LENGTH: 783

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)....(783)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-2039

Query Match 24.8%; Score 734.8; DB 10; Length 783;
Best Local Similarity 98.3%; Pred. No. 6.7e-222;
Matches 772; Conservative 0; Mismatches 10; Indels 3; Gaps 3;

Qy 141 CGGCGCGCAAGACCCGCTGGCGCACTTGCACACCGAGAGAGCGCGACCGCTGGGG 200
Db 1 CGGCGCGCAAGACCCGCTGGCGCACTTGCACACCGAGAGAGCGCGACCGCTGGGG 60

Qy 201 TGCTCGGCGGCCCAACACCGGTGTACCTGCGAGGTGGTGGCAGCGGTAGCGGACTCG 260
Db 61 TGCTCGGCGGCCCAACACCGGTGTACCTGCGAGGTGGTGGCAGCGGTA-CCGGGACTCG 119

Qy 261 GCGCGCGCGCTCTACGTCTTCTCCGAGTTCAACCGGTATCTTCAACTGTGGAGAGCG 320
Db 120 GCGCGCGCGCTCTACGTCTTCTCCGAGTTCAACCGGTATCTTCAACTGTGGAGAGCG 179

Qy 321 GTTCAGAGACTCATGCGAGGACCAAGTTAAAGTTGCTCGCTGGACCAACATATTCCTG 380
Db 180 GTTCAGAGACTCATGCGAGGACCAAGTTAAAGTTGCTCGCTGGACCAACATATTCCTG 239

Qy 381 ACACGAGTGCAGTGTCTAATGTTGGGCGTTAAGTGGATGATTTCTTACTTTAAAGGAA 440
Db 240 ACACGAGTGCAGTGTCTAATGTTGGGCGTTAAGTGGATGATTTCTTACTTTAAAGGAA 299

Qy 441 ACCGGCTTCCAAAGTGTGCTACTTCTGACCTCCACACTGGAATACTCCGAAGCA 500
Db 300 ACCGGCTTCCAAAGTGTGCTACTTCTGACCTCCACACTGGAATACTCCGAAGCA 359

Qy 501 ATCAAAATATTTCTGGTCCATGAAAGGAATAGAACTGGCTGTGGGCGCCCACTCTGCC 560
Db 360 ATCAAAATATTTCTGGTCCATGAAAGGAATAGAACTGGCTGTGGGCGCCCACTCTGCC 419

Qy 561 CCAGAATACGAGTGAACCATGACAGTTTACAGATCCCAATACACAGTGAACAGAGG 620
Db 420 CCAGAATACGAGTGAACCATGACAGTTTACAGATCCCAATACACAGTGAACAGAGG 479

Qy 621 AGGGAAACCAACCATGGAGAGTCCAGAAAGGCTCTCAGCAGGCTCAGTCCAGAG 680
Db 480 AGGGAAACCAACCATGGAGAGTCCAGAAAGGCTCTCAGCAGGCTCAGTCCAGAG 539

Qy 681 CGATCTTCAGACTCCGAG-TCGAATGAAATGAGCCACACCTTCCACATGGTGTAGCCA 739
Db 540 CGATCTTCAGACTCCGAGTTCGAATGAAATGAGCCACACCTTCCACATGGTGTAGCCA 599

Qy 740 GAGAGAGGGTCAAGGACTCTTCCCTGGTGTAGTCTTCACTGTAAAGTCACTTAA 799
Db 600 GAGAGAGGGTCAAGGACTCTTCCCTGGTGTAGTCTTCACTGTAAAGTCACTTAA 659

Qy 800 GAGAGAAACTTCTGGTCTCAAGCAAGGAGTGGGCTCCCACTTGGGACAGCTCG 859
Db 660 GAGAGAAACTTCTGGTCTCAAGCAAGGAGTGGGCTCCCACTTGGGAA-AACTCG 718

Qy 860 CATCGCTCCCATCATTTGCTGTCAAGGACGGGAAAGCATCACTCATGAAGGAGAGA 919
Db 719 CATCNCTCCCATCATTTGCTGTCAAGGACNGNAAACCAACCCATTAAGGAAAGA 778

Qy 920 GATTT 924
Db 779 GATTT 783

RESULT 12
US-09-988-626-28
; Sequence 28, Application US/09988626
; Publication No. US2003004959A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques

; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,626
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 26664
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (910)..(13104)
; OTHER INFORMATION: exon 1: 910-1154; exon 2: 1736-1786; exon 3:
; OTHER INFORMATION: 1925-1995; exon 4: 3025-3089; exon 5: 4361-4418;
; OTHER INFORMATION: exon 6: 5582-5650; exon 7: 7075-7194; exon 8:
; OTHER INFORMATION: 8186-8244; exon 9: 12878-12936; exon 10:
; OTHER INFORMATION: 13032-13104;
; NAME/KEY: misc feature
; LOCATION: (13756)..(22917)
; OTHER INFORMATION: exon 11: 13756-13868; exon 12: 15283-15378; exon
; OTHER INFORMATION: 13: 16278-16416; exon 14: 16494-16583; exon 15:
; OTHER INFORMATION: 18583-18701; exon 16: 20349-20445; exon 17:
; OTHER INFORMATION: 22172-22310; exon 18: 22879-22917
; NAME/KEY: misc feature
; LOCATION: (23045)..(26452)
; OTHER INFORMATION: exon 19: 23045-23154; exon 20: 23795-23895; exon
; OTHER INFORMATION: 21: 23973-24093; exon 22: 24354-24432; exon 23:
; OTHER INFORMATION: 25026-25170; exon 24: 25812-26036; polyadenylation
; OTHER INFORMATION: signal: 26447-26452
; NAME/KEY: variation
; LOCATION: (826)..(23879)
; OTHER INFORMATION: s at positions 826 and 23180 is G or C; y at
; OTHER INFORMATION: positions 1914, 568, 7165, 16431, 1857 and 20486
; OTHER INFORMATION: is C or T; n at position 13128 is t or tgat; r at
; OTHER INFORMATION: positions 22211 and 23879 is A or G.
US-09-988-626-28

Query Match 22.2%; Score 657.2; DB 9; Length 26664;
Best Local Similarity 99.5%; Pred. No. 2e-196;
Matches 659; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2297 CATGAAGTCTGCTTTGGAGACTTTTCCAAACATGCCAAGCTGATTTCCCCCACTGAAAGC 2356
Db 25805 CTTCTAGTCTGCTTTGGAGACTTTTCCAAACATGCCAAGCTGATTTCCCCCACTGAAAGC 25864

Qy 2357 CTTGTTTGTGCGCAGATCGAGGAGATGGAGAGCGCGGAGAGAGCGGAGCTTCGGCA 2416
Db 25865 CTTGTTTGTGCGCAGATCGAGGAGATGGAGAGCGCGGAGAGAGCGGAGCTTCGGCA 25924

Qy 2417 GGTGCGGCGGCTTCTGTCAGGAGCTGCGAGCGGCTTGGAGGATGGGAGCTTCA 2476
Db 25925 GGTGCGGCGGCTTCTGTCAGGAGCTGCGAGCGGCTTGGAGGATGGGAGCTTCA 25984

Qy 2477 GCAGAAGCGGCGCCACACAGAGAGCCACAGCCCAAGAGGTCACAGCCCACTGAGATC 2536
Db 25985 GCAGAAGCGGCGCCACACAGAGAGCCACAGCCCAAGAGGTCACAGCCCACTGAGATC 26044

Qy 2537 TGGAGAGCCCTGAACTCAGAAGCTGTGTCTTCTGCCCCACGACCGCCGATCTG 2596
Db 26045 TGGAGAGCCCTGAACTCAGAAGCTGTGTCTTCTGCCCCACGACCGCCGATCTG 26104

Qy 2597 CCTCTCTGCTGGTAGAGCTGAAGAGCAGCTCCCCAGGAGGAGCTCAGATAGGTG 2656
|||


```
Db 26105 CCCTCTTGGTGTAGAGCTGAAGACGACGCTCCCCCAGGAGGAGCTCAGGATAGGTG 26164
Qy 2657 GTATGGAGCTGTGCGAGGCTTGGGCTCCACATGAAGCACTAGTCTATAGATGCTCTTTA 2716
Db 26165 GTATGGAGCTGTGCGAGGCTTGGGCTCCACATGAAGCACTAGTCTATAGATGCTCTTTA 26224
Qy 2717 GGACTGTGCTTGGGCACAGCGCGGCGGCGGAGGCTGCCACACGGAAGCAAGCAGATGAA 2776
Db 26225 GGACTGTGCTTGGGCACAGCGCGGCGGCGGAGGCTGCCACACGGAAGCAAGCAGATGAA 26284
Qy 2777 CTAATTTTCATTTCAAGGAGTGTGTTTAAAGAGTCTTGGAAACAGACGGCGGCACCTTTCC 2836
Db 26285 CTAATTTTCATTTCAAGGAGTGTGTTTAAAGAGTCTTGGAAACAGACGGCGGCACCTTTCC 26344
Qy 2837 TCTAATCCAGCAAGTGTATCTCCTGCACACAGACAGACAGAGTAACAGATCAGTGG 2896
Db 26345 TCTAATCCAGCAAGTGTATCTCCTGCACACAGACAGACAGAGTAACAGATCAGTGG 26404
Qy 2897 GTCTAAGTGTCCGAGACTTAAGCAAAATAGTATTTTCAGCTGCAATAAAGATTTGCTTC 2956
Db 26405 GTCTAAGTGTCCGAGACTTAAGCAAAATAGTATTTTCAGCTGCAATAAAGATTTGCTTC 26464
Qy 2957 AA 2958
Db 26465 AA 26466

RESULT 13
US-09-988-687-28
; Sequence 28, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
; APPLICANT: Tavtignan, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,687
; PRIORITY FILING DATE: 2001-11-20
; PRIORITY APPLICATION NUMBER: 09/564,805
; PRIORITY FILING DATE: 2000-05-05
; PRIORITY APPLICATION NUMBER: US 60/107,468
; PRIORITY FILING DATE: 1998-11-06
; PRIORITY APPLICATION NUMBER: 09/434,382
; PRIORITY FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 26664
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (910)..(13104)
; OTHER INFORMATION: exon 1: 910-1154; exon 2: 1736-1786; exon 3:
; OTHER INFORMATION: 1925-1995; exon 4: 3025-3089; exon 5: 4361-4418;
; OTHER INFORMATION: exon 6: 5582-5650; exon 7: 7075-7194; exon 8:
; OTHER INFORMATION: 8186-8244; exon 9: 12878-12936; exon 10:
; OTHER INFORMATION: 13032-13104;
; NAME/KEY: misc_feature
; LOCATION: (13756)..(22917)
; OTHER INFORMATION: exon 11: 13756-13868; exon 12: 15283-15378; exon
; OTHER INFORMATION: 13: 16278-16416; exon 14: 16498-16583; exon 15:
; OTHER INFORMATION: 18583-18701; exon 16: 20349-20445; exon 17:
; OTHER INFORMATION: 22172-22310; exon 18: 22879-22917
; NAME/KEY: misc_feature
; LOCATION: (23045)..(26452)
; OTHER INFORMATION: exon 19: 23045-23154; exon 20: 23795-23895; exon
; OTHER INFORMATION: 21: 23973-24093; exon 22: 24354-24432; exon 23:
; OTHER INFORMATION: 25026-25170; exon 24: 25812-26036; polyadenylation
; * TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
```

```
; OTHER INFORMATION: signal: 26447-26452
; NAME/KEY: variation
; LOCATION: (826)..(23879)
; OTHER INFORMATION: s at positions 826 and 23180 is G or C; y at
; OTHER INFORMATION: positions 1914, 5568, 7165, 16431, 1857 and 20486
; OTHER INFORMATION: is C or T; n at position 13128 is t or tgat; r at
; OTHER INFORMATION: positions 22211 and 23879 is A or G.
US-09-988-687-28

Query Match 22.2%; Score 657.2; DB 9; Length 26664;
Best Local Similarity 99.5%; Pred. No. 2e-196;
Matches 659; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2297 CATGAAGTCTGCTTTGGAGACTTTCCAAACAATGCCCAAGCTGATTTCCCCACCTGAAGC 2356
Db 25805 CTTCTAGGCTGTGCTTTGGAGACTTTCCAAACAATGCCCAAGCTGATTTCCCCACCTGAAGC 25864
Qy 2357 CCTCTTTGCTGGCGACATCGAGGAGATGGAGGAGCGCAGGAGAAAGCGGAGCTGCCGCA 2416
Db 25865 CTTGTTTGTCTGGCGACATCGAGGAGATGGAGGAGCGCAGGAGAAAGCGGAGCTGCCGCA 25924
Qy 2417 GGTGCGGCGCGCCCTCCTGTCTCAGGGAGCTGGCAGCGCGCTTGGAGATGGGGAGCCTCA 2476
Db 25925 GGTGCGGCGCGCCCTCCTGTCTCAGGGAGCTGGCAGCGCGCTTGGAGATGGGGAGCCTCA 25984
Qy 2477 GCAGAAGCGGGCCACACAGAGGAGCCACAGGCCAAGAAGTCAAGGCCCAAGTGAAGATC 2536
Db 25985 GCAGAAGCGGGCCACACAGAGGAGCCACAGGCCAAGAAGTCAAGGCCCAAGTGAAGATC 26044
Qy 2537 TGGGAGACCTGAACTCAGAAAGCTGTGTCTTCTCCCCACGACGACCCCTATCTG 2596
Db 26045 TGGGAGACCTGAACTCAGAAAGCTGTGTCTTCTCCCCACGACGACCCCTATCTG 26104
Qy 2597 CCCTCCTTGTCTGTAGAACTGAAGAGCACGCTGCCCCAGGAGGACGCTCAGGATAGGTG 2656
Db 26105 CCCTCCTTGTCTGTAGAACTGAAGAGCACGCTGCCCCAGGAGGACGCTCAGGATAGGTG 26164
Qy 2657 GTATGGAGCTGTGCGGAGGCTTGGGCTCCACATGAAGCACTAGTCTATAGATGCTCTTA 2716
Db 26165 GTATGGAGCTGTGCGGAGGCTTGGGCTCCACATGAAGCACTAGTCTATAGATGCTCTTA 26224
Qy 2717 GGACTGTGCTTGGCACAGCGCGGCGGCGGAGGCTGCCACACGGAAGCAGATGAA 2776
Db 26225 GGACTGTGCTTGGCACAGCGCGGCGGCGGAGGCTGCCACACGGAAGCAGATGAA 26284
Qy 2777 CTAATTTTCATTTCAAGGAGTGTGTTTAAAGAGTCTTGGAAACAGACGGCGGCACCTTTCC 2836
Db 26285 CTAATTTTCATTTCAAGGAGTGTGTTTAAAGAGTCTTGGAAACAGACGGCGGCACCTTTCC 26344
Qy 2837 TCTAATCCAGCAAGTGTATCTCCTGCACACAGAGAGCAAGCAGATGATGAGTGG 2896
Db 26345 TCTAATCCAGCAAGTGTATCTCCTGCACACAGAGAGCAAGCAGATGATGAGTGG 26404
Qy 2897 GTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTTCAGCTGCAATAAAGATTTGCTTC 2956
Db 26405 GTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTTCAGCTGCAATAAAGATTTGCTTC 26464
Qy 2957 AA 2958
Db 26465 AA 26466

RESULT 14
US-09-988-626-27
; Sequence 27, Application US/09988626
; Publication No. US20030044959A1
; GENERAL INFORMATION:
; APPLICANT: Tavtignan, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
```


;; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
;; FILE REFERENCE: 2318-258
;; CURRENT APPLICATION NUMBER: US/09/988,626
;; PRIORITY FILING DATE: 2001-11-20
;; PRIOR APPLICATION NUMBER: 09/564,805
;; PRIOR FILING DATE: 2000-05-05
;; PRIOR APPLICATION NUMBER: US 60/107,468
;; PRIOR FILING DATE: 1998-11-06
;; PRIOR APPLICATION NUMBER: 09/434,382
;; PRIOR FILING DATE: 1999-11-05
;; NUMBER OF SEQ ID NOS: 240
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 27
;; LENGTH: 655
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (1)..(228)
;; OTHER INFORMATION: exon 24
;; NAME/KEY: polyA signal
;; LOCATION: (636)..(641)
US-09-988-626-27

Query Match 22.1%; Score 655; DB 9; Length 655;
Best Local Similarity 100.0%; Pred. No. 1.1e-196;
Matches 655; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2304 GTCTGCTTTGGAGACTTCCAAACATGCCAAGCTGATTTCCCTCCCTGAAAGCCCTGTTT 2363
Db 1 GTCTGCTTTGGAGACTTCCAAACATGCCAAGCTGATTTCCCTCCCTGAAAGCCCTGTTT 60
QY 2364 GTTGGCGACATCGAGGAGATGGAGGAGCGCAGGAGAGCGGAGCTGCGGAGGTGCGG 2423
Db 61 GCTGGCGACATCGAGGAGATGGAGGAGCGCAGGAGAGCGGAGCTGCGGAGGTGCGG 120
QY 2424 GGGGCCCTCTGTCCAGGAGCTGGCAGGGGCTGGAGAGTGGGAGGCTCAGCAGAG 2483
Db 121 GGGGCCCTCTGTCCAGGAGCTGGCAGGGGCTGGAGAGTGGGAGGCTCAGCAGAG 180
QY 2484 CGGGCCACACAGAGGAGCCACAGGCGCAAGAGGTGAGAGCCAGTGAAGATCTGGGAGA 2543
Db 181 CGGGCCACACAGAGGAGCCACAGGCGCAAGAGGTGAGAGCCAGTGAAGATCTGGGAGA 240
QY 2544 CCCTGAATCTCAGAAGGCTGTGTCTTCTGCCCCACGACCGACCCGATCTGCCCTCT 2603
Db 241 CCCTGAATCTCAGAAGGCTGTGTCTTCTGCCCCACGACCGACCCGATCTGCCCTCT 300
QY 2604 TGTGTGAGAGCTGAAGAGCGCTGCTCCCTCCAGGAGCGAGCTCAGGATAGGTGGTATGA 2663
Db 301 TGTGTGAGAGCTGAAGAGCGCTGCTCCCTCCAGGAGCGAGCTCAGGATAGGTGGTATGA 360
QY 2664 GTGTGCCAGGCTGGGCTCCACATAGCACTAGTCTATAGATGCCTCTTAGGACTGG 2723
Db 361 GTGTGCCAGGCTGGGCTCCACATAGCACTAGTCTATAGATGCCTCTTAGGACTGG 420
QY 2724 TGCCTGGCAGCGCGGCGCAGGAGGCTGCCACACGGAAGCAAGCAGATGAATTAATTT 2783
Db 421 TGCCTGGCAGCGCGGCGCAGGAGGCTGCCACACGGAAGCAAGCAGATGAATTAATTT 480
QY 2784 CATTTCAAGGAGCTTTTAAAGAGTCTTTGGAAACAGACGGGGCACCTTTCTCTAATC 2843
Db 481 CATTTCAAGGAGCTTTTAAAGAGTCTTTGGAAACAGACGGGGCACCTTTCTCTAATC 540
QY 2844 CAGCAAGTGTATTCCTGACACACAGAGCAGCAGATGAACAGGATCAGTGGTCTAAG 2903
Db 541 CAGCAAGTGTATTCCTGACACACAGAGCAGCAGATGAACAGGATCAGTGGTCTAAG 600
QY 2904 TGTCCGAGACTTTAAACGAAATAGTATTTTTCAGCTGCAATTAAGATTGAGTTGCAA 2958
Db 601 TGTCCGAGACTTTAAACGAAATAGTATTTTTCAGCTGCAATTAAGATTGAGTTGCAA 655

RESULT 15
US-09-988-687-27
;; Sequence 27, Application US/09988687
;; Publication No. US20030045704A1
;; GENERAL INFORMATION:
;; APPLICANT: Tavtigian, Sean V.
;; APPLICANT: Teng, David H.F.
;; APPLICANT: Simard, Jacques
;; APPLICANT: Rommens, Johanna M.
;; APPLICANT: Myriad Genetics, Inc.
;; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
;; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
;; FILE REFERENCE: 2318-258
;; CURRENT APPLICATION NUMBER: US/09/988,687
;; PRIORITY FILING DATE: 2001-11-20
;; PRIOR APPLICATION NUMBER: 09/564,805
;; PRIOR FILING DATE: 2000-05-05
;; PRIOR APPLICATION NUMBER: US 60/107,468
;; PRIOR FILING DATE: 1998-11-06
;; PRIOR APPLICATION NUMBER: 09/434,382
;; PRIOR FILING DATE: 1999-11-05
;; NUMBER OF SEQ ID NOS: 240
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 27
;; LENGTH: 655
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (1)..(228)
;; OTHER INFORMATION: exon 24
;; NAME/KEY: polyA signal
;; LOCATION: (636)..(641)
US-09-988-687-27

Query Match 22.1%; Score 655; DB 9; Length 655;
Best Local Similarity 100.0%; Pred. No. 1.1e-196;
Matches 655; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2304 GTCTGCTTTGGAGACTTCCAAACATGCCAAGCTGATTTCCCTCCCTGAAAGCCCTGTTT 2363
Db 1 GTCTGCTTTGGAGACTTCCAAACATGCCAAGCTGATTTCCCTCCCTGAAAGCCCTGTTT 60
QY 2364 GCTGGCGACATCGAGGAGATGGAGGAGCGCAGGAGAGCGGAGCTGCGGAGGTGCGG 2423
Db 61 GCTGGCGACATCGAGGAGATGGAGGAGCGCAGGAGAGCGGAGCTGCGGAGGTGCGG 120
QY 2424 GGGGCCCTCTGTCCAGGAGCTGGCAGGGGCTGGAGAGTGGGAGGCTCAGCAGAG 2483
Db 121 GGGGCCCTCTGTCCAGGAGCTGGCAGGGGCTGGAGAGTGGGAGGCTCAGCAGAG 180
QY 2484 CGGGCCACACAGAGGAGCCACAGGCGCAAGAGGTGAGAGCCAGTGAAGATCTGGGAGA 2543
Db 181 CGGGCCACACAGAGGAGCCACAGGCGCAAGAGGTGAGAGCCAGTGAAGATCTGGGAGA 240
QY 2544 CCCTGAATCTCAGAAGGCTGTGTCTTCTGCCCCACGACCGACCCGATCTGCCCTCT 2603
Db 241 CCCTGAATCTCAGAAGGCTGTGTCTTCTGCCCCACGACCGACCCGATCTGCCCTCT 300
QY 2604 TGTGTGAGAGCTGAAGAGCGCTGCTCCCTCCAGGAGCGAGCTCAGGATAGGTGGTATGA 2663
Db 301 TGTGTGAGAGCTGAAGAGCGCTGCTCCCTCCAGGAGCGAGCTCAGGATAGGTGGTATGA 360
QY 2664 GTGTGCCAGGCTGGGCTCCACATAGCACTAGTCTATAGATGCCTCTTAGGACTGG 2723
Db 361 GTGTGCCAGGCTGGGCTCCACATAGCACTAGTCTATAGATGCCTCTTAGGACTGG 420
QY 2724 TGCCTGGCAGCGCGGCGCAGGAGGCTGCCACACGGAAGCAAGCAGATGAATTAATTT 2783
Db 421 TGCCTGGCAGCGCGGCGCAGGAGGCTGCCACACGGAAGCAAGCAGATGAATTAATTT 480
QY 2784 CATTTCAAGGAGCTTTTAAAGAGTCTTTGGAAACAGACGGGGCACCTTTCTCTAATC 2843
Db 481 CATTTCAAGGAGCTTTTAAAGAGTCTTTGGAAACAGACGGGGCACCTTTCTCTAATC 655

Db 481 CATTTCAGGCAGTTTTTAAAGAAGCTTTGGAACAGACGCGGCACCTTTCCTCTAATC 540
Qy 2844 CAGCAAGTGATTCCTGCACACAGAGACAGAGTAACAGGATCAGTGGGTCTAAG 2903
Db 541 CAGCAAGTGATTCCTGCACACAGAGACAGAGTAACAGGATCAGTGGGTCTAAG 600
Qy 2904 TGTCCGAGACTTAACGAAAATAGTATTTTCAGCTGCAATAAAGATTGAGTTTGCAA 2958
Db 601 TGTCCGAGACTTAACGAAAATAGTATTTTCAGCTGCAATAAAGATTGAGTTTGCAA 655

Search completed: May 18, 2003, 02:19:09
Job time : 522 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: May 17, 2003, 15:56:54 ; Search time 6428.12 Seconds
(without alignments)
11232.524 Million cell updates/sec

Title: US-09-434-382-1
Perfect score: 2481
Sequence: 1 atgtggcgctttgctgcgt.....agaaggtcagagcccaagtga 2481

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2481	100.0	2908	9	AF304370	AF304370 Homo sapi
2	2481	100.0	2992	6	AX405825	AX405825 Sequence
3	2477.8	99.9	2997	9	BC001939	BC001939 Homo sapi
4	2477.8	99.9	3006	9	BC004158	BC004158 Homo sapi
5	2455.4	99.0	2908	9	AF308698	AF308698 Pan trogl
6	2448.2	98.7	2976	9	AK001392	AK001392 Homo sapi
7	2442.6	98.5	2893	9	AF308694	AF308694 Gorilla g
8	2349	94.7	2907	9	AK094687	AK094687 Homo sapi
9	1745	70.3	2734	9	AK094333	AK094333 Homo sapi
10	1715.2	69.1	2163	9	AK074244	AK074244 Homo sapi
11	1686.2	68.0	2712	10	AF308696	AF308696 Mus muscu
12	823	33.2	2793	9	AK094012	AK094012 Homo sapi
13	247.4	10.0	118788	9	AC005277	AC005277 Homo sapi
14	156.8	6.3	679	9	HS329331	HS329331 Homo sapi
15	146.2	5.9	740	9	AF304371S2	AF304369 Homo sapi
16	144.4	5.8	34593	10	AF348157	AF348157 Mus muscu
17	144.4	5.8	196566	10	AL663045	AL663045 Mouse DNA
C 18	135.8	5.5	167336	2	AC128361	AC128361 Rattus no
C 19	135.8	5.5	186836	2	AC097911	AC097911 Rattus no
C 20	133.4	5.4	2690	3	AF215894	AF215894 Drosophil
21	133.4	5.4	2728	3	AV119279	AV119279 Drosophil
22	133.4	5.4	115873	2	AC017383	AC017383 Drosophil
C 23	133.4	5.4	175681	3	AC007417	AC007417 Drosophil
24	133.4	5.4	190574	2	AC007352	AC007352 Drosophil
C 25	133.4	5.4	240481	3	AE003830	AE003830 Drosophil
C 26	132.4	5.3	186836	2	AC097911	AC097911 Rattus no
27	124	5.0	721	9	AF304371S1	AF304371 Homo sapi
28	115.2	4.6	167336	2	AC128361	AC128361 Rattus no
29	106.6	4.3	2704	8	AY117303	AY117303 Arabidops
C 30	95.4	3.8	2818	8	AY070061	AY070061 Arabidops
C 31	95.4	3.8	752	9	HS329367	HS329367 Homo sapi
C 32	90.2	3.6	1617	8	AY086950	AY086950 Arabidops
C 33	84.2	3.4	31201	8	SPAC1D4	Z69239 S.pombe chr
C 34	71.8	2.9	46335	8	AC022354	AC022354 Arabidops
C 35	70	2.8	19158	8	AF188714	AF188714 Emericell
C 36	67	2.7	32329	8	SPBC3D6	Z95620 S.pombe chr
C 37	65	2.6	143209	8	AF001550	AF001550 Oryza sat
C 38	63.4	2.6	125020	9	AF429315	AF429315 Homo sapi
C 39	59.4	2.4	256050	1	AL627274	AL627274 Salmonell
40	58.8	2.4	75289	8	AB023046	AB023046 Arabidops
41	58.4	2.4	2517	6	AR137565	AR137565 Sequence
42	58.4	2.4	2517	6	AR139071	AR139071 Sequence
C 43	58.4	2.4	3685	8	SCYKR079C	Z28304 S.cerevisia
44	57.8	2.3	23630	1	AE008803	AE008803 Salmonell
45	56.8	2.3	23533	1	AE001039	AE001039 Archaeogl

ALIGNMENTS

RESULT 1
AF304370
LOCUS
DEFINITION Homo sapiens putative prostate cancer susceptibility protein
ACCESSION AF304370
VERSION AF304370.1
KEYWORDS HPC2/ELAC2 mRNA, complete cds.
SOURCE AF304370.1 GI:10880932
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2908)
REFERENCE
Tavtigian,S.V., Simard,J., Teng,D.H.F., Abtin,V., Baumgard,M.,
Beck,A., Camp,N.J., Carillo,A.R., Chen,Y., Dayananth,P.,
2908 bp mRNA linear PRI 23-FEB-2001

Desrochers, M., Dumont, M., Farnham, J. M., Frank, D., Frye, C.,
Gaffari, S., Gupta, J. S., Hu, R., Iliev, D., Janeski, T., Kort, E. N.,
Laity, K. E., Leavitt, A., Leblanc, G., McArthur-Morrison, J.,
Pederson, A., Penn, B., Peterson, K. T., Reid, J. E., Richards, S.,
Schroeder, M., Smith, R., Snyder, S. C., Swedlund, B., Swensen, J.,
Thomas, A., Tranchant, M., Woodland, A. M., Labrie, F., Skolnick, M. H.,
Neuhausen, S., Rommens, J. and Cannon-Albright, L. A.
A candidate prostate cancer susceptibility gene at chromosome 17p
Nat. Genet. 27 (2), 172-180 (2001)
21096977
11175785
2 (bases 1 to 2908)
Tavtigian, S. V., Simard, J., Teng, D. H. F., Baumgard, M., Beck, A.,
Camp, N. J., Carillo, A. R., Chen, Y., Dayananth, P., Desrochers, M.,
Dumont, M., Farnham, J. M., Frank, D., Frye, C., Gaffari, S.,
Gupta, J. S., Hu, R., Iliev, D., Janeski, T., Kort, E. N., Laity, K. E.,
Leavitt, A., Leblanc, G., McArthur-Morrison, J., Pederson, A., Penn, B.,
Peterson, K. T., Reid, J. E., Richards, S., Schroeder, M., Smith, R.,
Snyder, S. C., Swedlund, B., Swensen, J., Thomas, A., Tranchant, M.,
Woodland, A. M., Labrie, F., Skolnick, M. H., Neuhausen, S., Rommens, J.,
and Cannon-Albright, L. A.
Direct Submission
Submitted (12-SEP-2000) Myriad Genetics, Inc., 320 Wakata Way, Salt
Lake City, UT 84108, USA
Location/Qualifiers
1..2908
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17p"
1..2481
/codon_start=1
/product="putative prostate cancer susceptibility protein
HPC2/ELAC2"
/protein_id="AAG24441.1"
/db_xref="GI:10880933"
/translation="MWALCSLLRSAAAGTWSQRTISQAPARRERPKDPLRLHRTPE
KRGFGSGGNTVYLVQVAAAGSDGAALVYFSEFNRYLFCNGEGVQRLWQERLKV
ARLDNIIFLTRHNSVNGSLSMILTKETGLPKVLSGPPQLEKYLEIAKIFSGPLKG
IELAVRPHSAPEDEMTYVQIPIHSEQRGRHQPWPSPRLSPRSDDSEN
ENEPHLPHGVSRQGRVDSLLVAFICKLHLKRGNFVLVAKEMGLPVGTAALAPIA
AVDKGKITHREGLIAELCTPPDPGAAFFVVECPDESFIQPCENATFORYOGKAD
APVALVHMAPASVLDVSRVVOOMERFGPDTHLVNLNENCAVHNLSFKIQTQNLNI
HPDIPFLTSFRCCKEPTLSVPMVQSECLLKYLQRRRWRDRIITCNPFEIVEA
LQPLFQSVQVEYRRSQDGPAPAKRSQYPIIFLGTGSAIPMKIRNVSAIIVNISP
DTSLLDCGEGTFOGLCRHYGDQDVRVLGTLAAVFSHLHADHTGTPSILLQERAL
ASLGLPLHLVAVNQLKWLQOYHNCQCEVLHHISMIPAKCLQEGAEISSPAVERL
ISSLRKDLPEEFOTCLVRHCKHAFGCALVHTSGWKVYSGDTWPCALVRMGKDATL
LIHEATLEDGLEEAETKSTTSQALSVGMRMNAEFIMLNHFSQRYAKVPLSPNFS
EKVGVADHMKVCGDPTPDKLIPPLKALFAGLIEBERERREKRELRQVRAALLSRE
LAGGLEDEPQQRAHTEBPQAKKVRQA"

BASE COUNT 702 a 793 c 826 g 587 t

ORIGIN

Query Match 100.0%; Score 2481; DB 9; Length 2908;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ATGTGGCGCTTTGCTCGTGTGCGGTCGCGGCGCGGACGACCATGTGCGAGGAGCGC 60
Db 1 ATGTGGCGCTTTGCTCGTGTGCGGTCGCGGCGCGGACGACCATGTGCGAGGAGCGC 60
Oy 61 ACCATATCGAGGACCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
Db 61 ACCATATCGAGGACCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
Oy 121 CGCAGCGAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
Db 121 CGCAGCGAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
Oy 181 CAGGTGGTGGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
Db 181 CAGGTGGTGGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240

```
QY 1321 TGCATCTCTGAGGAATTCATAGTTGAGGGCTGCGAGCTTCCCAACTTCCAGCAGAGCGTG 1380
Db 1321 TGCATCTCTGAGGAATTCATAGTTGAGGGCTGCGAGCTTCCCAACTTCCAGCAGAGCGTG 1380
QY 1381 CAGGAGTACAGGAGAGTGCAGCAGCAGCGCCAGCCCGAGCAGCAGAGAGAGAGAGTAC 1440
Db 1381 CAGGAGTACAGGAGAGTGCAGCAGCAGCGCCAGCCCGAGCAGCAGAGAGAGAGTAC 1440
QY 1441 CCAGAAATCATCTTCTTGGAACAGGGTCTGCCATCCCGATCCCGATCCGAAATGTCAAT 1500
Db 1441 CCAGAAATCATCTTCTTGGAACAGGGTCTGCCATCCCGATCCCGATCCGAAATGTCAAT 1500
QY 1501 GCCACACTTGTCAACATAGCCCGACACAGTCTCTGCTACTGAGTGTGAGGAGCACA 1560
Db 1501 GCCACACTTGTCAACATAGCCCGACACAGTCTCTGCTACTGAGTGTGAGGAGCACA 1560
QY 1561 TTTGGGCGAGCTGTGGCTCATTCAGGAGACAGGCTGGACAGGCTCTGGGACCCCTGGCT 1620
Db 1561 TTTGGGCGAGCTGTGGCTCATTCAGGAGACAGGCTGGACAGGCTCTGGGACCCCTGGCT 1620
QY 1621 GCTGTGTTGTGTCCTGACCTGACGAGATCACCACAGGGCTTCCCAAGTATCTTGCTG 1680
Db 1621 GCTGTGTTGTGTCCTGACCTGACGAGATCACCACAGGGCTTCCCAAGTATCTTGCTG 1680
QY 1681 CAGAGAGACGCGCTTGGCATCTTTGGAAAGCGCTTACCCCTTGTGCTGTTGCTGCC 1740
Db 1681 CAGAGAGACGCGCTTGGCATCTTTGGAAAGCGCTTACCCCTTGTGCTGTTGCTGCC 1740
QY 1741 CCCAACCCAGCTCAAGCCTGGCTCCAGCAGTACCAACACAGTGCAGGAGTCTGCGAC 1800
Db 1741 CCCAACCCAGCTCAAGCCTGGCTCCAGCAGTACCAACACAGTGCAGGAGTCTGCGAC 1800
QY 1801 CACATCAGTATGATTCCTGCCAAATGCCCTTCAGGAGGGGCTGAGATCTCCAGTCTCGCA 1860
Db 1801 CACATCAGTATGATTCCTGCCAAATGCCCTTCAGGAGGGGCTGAGATCTCCAGTCTCGCA 1860
QY 1861 GTGGAAGATGATGATGCTGCTGTCGAAATGATGATGATGATGATGATGATGATGATGAT 1920
Db 1861 GTGGAAGATGATGATGCTGCTGTCGAAATGATGATGATGATGATGATGATGATGATGAT 1920
QY 1921 CTGGTGGGCTGCAAGCAGTGGCTGCTGGCTGCTGGCTGCTGGCTGCTGGCTGCTGGCTG 1980
Db 1921 CTGGTGGGCTGCAAGCAGTGGCTGCTGGCTGCTGGCTGCTGGCTGCTGGCTGCTGGCTG 1980
QY 1981 GTGGTCTATTCCGGGGACACCATGCGCTCGAGGCTCTGGTCCGATGGGGAAGATGCC 2040
Db 1981 GTGGTCTATTCCGGGGACACCATGCGCTCGAGGCTCTGGTCCGATGGGGAAGATGCC 2040
QY 2041 ACCCTCTGATACATGAAGCCACCTTGGAAAGTGGTTTGGAAAGAGACATGGAAGAG 2100
Db 2041 ACCCTCTGATACATGAAGCCACCTTGGAAAGTGGTTTGGAAAGAGACATGGAAGAG 2100
QY 2101 ACACAGCAGACAGCTCCCAAGCCATCAGCTGGGATCGGATGACCGAGTTCATT 2160
Db 2101 ACACAGCAGACAGCTCCCAAGCCATCAGCTGGGATCGGATGACCGAGTTCATT 2160
QY 2161 ATGCTGAACCACTTCAGCCAGCGCTATGCAAGGTCCTCTTTCAGCCCCCACTTCAGC 2220
Db 2161 ATGCTGAACCACTTCAGCCAGCGCTATGCAAGGTCCTCTTTCAGCCCCCACTTCAGC 2220
QY 2221 GAGAAAGTGGGAGTTCCTTTGACCAATGAAGTGTCTGTTGGAGTTCCTTTCAGCAATG 2280
Db 2221 GAGAAAGTGGGAGTTCCTTTGACCAATGAAGTGTCTGTTGGAGTTCCTTTCAGCAATG 2280
QY 2281 CCCAGCTGATTTCCCACTGAAGCCCTGTTGCTGGGACATCGAGAGATGGAGGAG 2340
Db 2281 CCCAGCTGATTTCCCACTGAAGCCCTGTTGCTGGGACATCGAGAGATGGAGGAG 2340
QY 2341 CGCAGGAGAAAGCGGAGTTCGCGCAGGTCGCGGCGGCTCTCTGTCAGGAGTTCGCA 2400
Db 2341 CGCAGGAGAAAGCGGAGTTCGCGCAGGTCGCGGCGGCTCTCTGTCAGGAGTTCGCA 2400
QY 2401 GCGGCGCTGGAGATGGGAGCTTCAGCAGAAAGCGGCGGCCACACAGAGGAGCCAGGCG 2460
```

```
Db 2401 GCGGCGCTGGAGATGGGAGCTTCAGCAGAAAGCGGCGGCCACACAGAGAGCCAGGCG 2460
QY 2461 AAGAAGGTTCAGAGCCCACTGA 2481
Db 2461 AAGAAGGTTCAGAGCCCACTGA 2481

RESULT 2
AX405825
LOCUS AX405825
DEFINITION Sequence 240 from Patent WO222660.
ACCESSION AX405825
VERSION AX405825.1 GI:21439089
KEYWORDS SOURCE
ORGANISM human.
REFERENCE
AUTHORS Tang, Y.T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q.A., Ren, F.,
Xue, A.J., Yang, Y., Wehrman, T. and Drmanac, R.T.
TITLE Novel nucleic acids and polypeptides
JOURNAL Patent: WO 022660-A 240 21-MAR-2002;
HYSEQ, INC. (US)
FEATURES
Location/Qualifiers
source 1..2992
/organism="Homo sapiens"
/db_xref="taxon:9606"
75..2555
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD34932.1"
/db_xref="GI:21439090"
/translation="MMALCSILRSACRTWSQGTISOAPARRERPKDPLRLHRLTRE
KRSGSGSGPNTVYLVVAAGSDSAAALVVFSEPNRYLFCGEGVORLQWHEHLKV
ARLDNIETRMHNSVGLSMILTKETGLKCVLSPQLEKYLEKIFSGPLKG
IEAVRPHSAPEYEDMTVYQIFHSEORRKHOPWSPERPISLSPSPSSSEN
ENEPHPLHVSQRRVDSLVAFIKLHLKRGFLVLAKEMLGPGVTAIAPIIA
AVKDKSITHREILABELCTPPDGAAFVVECPDESFIQICENATFORYGKAD
APVALVHMAPASVLVDSRYCOOMEREGPDTQHLVLNENCASVNLRSKLTQTLNLI
HPDIFPLLTSTFRCKEGPTLSVPMVQSECLLKYLRRRERQORDAITCNBEETVEA
LDLPNFOQSVQERYRSADGPAPAEKRSQYPIIFLGTGSAIPMKIRNVSAITLVNIS
TDLLLDCGEGTGGQLCRHYGDQVDRVLGTAAVFSHLHADHTGTPSILLQERAL
ASLGLPLHLLVWPNQKALQYHNCQCEVLHHISMIPAKLCQEGAEISSPAVERL
ISSLLRTCDLEEFOTCLVRKCHAFGALVHTSGWKVYSGDTPWCEALVRMGKDATL
LIHETATLEDGLEEAEVETHTTQSAISVGMWNAEFLMLNHFORSYAKVPLFSPNFS
EKGVAFDHMKVCFEGDEPTMPKLIPLPKALFAGDIEEMEERERKRELRQVRAALLSRE
LAGGLEDBGPQQKRAHTEEPQAKVRAQ"
BASE COUNT 725 a 807 c 859 g 601 t
ORIGIN

Query Match 100.0%; Score 2481; DB 6; Length 2992;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTGGGCGCTTTGCTGCTGCTGCTGCGGCGCGGACGACCATGTCGCGAGGAGCGC 60
Db 75 ATGTGGGCGCTTTGCTGCTGCTGCTGCGGCGCGGACGACCATGTCGCGAGGAGCGC 134
QY 61 ACCATATCGCAGGACACCGCGCGCGCGAGCGCGCGCGAAGGACCCGCTGCGGACCTG 120
Db 135 ACCATATCGCAGGACACCGCGCGCGCGAGCGCGCGCGAAGGACCCGCTGCGGACCTG 194
QY 121 CGCAGCGCAGAGAGCGCGGACCGCTGCGGGTGTCTCCGGCGGCCCAACACCGTGTACCTG 180
Db 195 CGCAGCGCAGAGAGCGCGGACCGCTGCGGGTGTCTCCGGCGGCCCAACACCGTGTACCTG 254
QY 181 CAGGTGTGGCAGCGGGTAGCCGGACTCGGCGCGCGCTCTACGTCTTCTCCAGTTC 240
Db 255 CAGGTGTGGCAGCGGGTAGCCGGACTCGGCGCGCGCTCTACGTCTTCTCCAGTTC 314
QY 241 AACCGGTATCTTTCAACTGTGTGAGAGCGGTTTCAGAGACTCATGAGAGCACAAGTTTA 300
```

Db 315 |||||AACCGGTATCTCTTCAA|||CTGTGGAGAGCGCTT|||CAGAGACT|||CATGCAGAGACACAAGTTA||| 374
Qy 301 AAGGTTGCTCGCCTGGGACAAATATCTCTGACACAAATGCACTGTGCTTAATGTTGGGGGC 360
Db 375 AAGGTTGCTCGCCTGGGACAAATATCTCTGACACAAATGCACTGTGCTTAATGTTGGGGGC 434
Qy 361 TTAAGTGAATGATCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTAATTTCTCGA 420
Db 435 TTAAGTGAATGATCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTAATTTCTCGA 494
Qy 421 CCTCCACAACCTGGAAAAATACCTCGAAGCAATCAAAATATTTCTGTGCCATTGAAAGGA 480
Db 495 CCTCCACAACCTGGAAAAATACCTCGAAGCAATCAAAATATTTCTGTGCCATTGAAAGGA 554
Qy 481 ATAGAACTGGCTGTCGGCCCCACTCTCTCCCCAGAAATACGAGGATGAAACCATGACAGTT 540
Db 555 ATAGAACTGGCTGTCGGCCCCACTCTCTCCCCAGAAATACGAGGATGAAACCATGACAGTT 614
Qy 541 TACAGATCCCCATACACAGTGAAACAGAGAGGGGAAAGCACCAACCATGGCAGAGTCCA 600
Db 615 TACAGATCCCCATACACAGTGAAACAGAGAGGGGAAAGCACCAACCATGGCAGAGTCCA 674
Qy 601 GAAAGGCTCTCAGCAGGCTCAGTCAGAGCGATCTTCTCAGACTCCGAGTCCGAATGAAAT 660
Db 675 GAAAGGCTCTCAGCAGGCTCAGTCAGAGCGATCTTCTCAGACTCCGAGTCCGAATGAAAT 734
Qy 661 GAGCACACCTTCCATCTGTTGTTAGCCAGAGAAAGGGGTGAGGACTCTTCTCCCTGTGC 720
Db 735 GAGCCACACCTTCCATCTGTTGTTAGCCAGAGAAAGGGGTGAGGACTCTTCTCCCTGTGC 794
Qy 721 GTAGCTTTTATCTGTAAGTTTCACTTTAAGAGAGAAACTTCTTGGTGTCTCAAGCAAG 780
Db 795 GTAGCTTTTATCTGTAAGTTTCACTTTAAGAGAGAAACTTCTTGGTGTCTCAAGCAAG 854
Qy 781 GAGATGGGCTCCCAAGTTGGGACAGCTGCCATCGCTCCCATCATTTGCTGTCTCAAGAC 840
Db 855 GAGATGGGCTCCCAAGTTGGGACAGCTGCCATCGCTCCCATCATTTGCTGTCTCAAGAC 914
Qy 841 GGGAAAGCATCACTCATGAAGGAAGAGATTTTGGCTGAAGAGCTGTGTACTCCTCCA 900
Db 915 GGGAAAGCATCACTCATGAAGGAAGAGATTTTGGCTGAAGAGCTGTGTACTCCTCCA 974
Qy 901 GATCTGTGTGCTTTTGTGTGTGTAGAACTCTCAGATGAAGCTTCAATCAACCCATC 960
Db 975 GATCTGTGTGCTTTTGTGTGTGTAGAACTCTCAGATGAAGCTTCAATCAACCCATC 1034
Qy 961 TGTGAAATGCCACCTTTTCAAGAGTACCAAGGAAAGGAGATGCCCGCTGGCTTGGTG 1020
Db 1035 TGTGAAATGCCACCTTTTCAAGAGTACCAAGGAAAGGAGATGCCCGCTGGCTTGGTG 1094
Qy 1021 GTTCATGTGGCCCCAGCATCTGTGTTGTGGACACAGGTACCCAGCAGTGGATGAGAGG 1080
Db 1095 GTTCATGTGGCCCCAGCATCTGTGTTGTGGACACAGGTACCCAGCAGTGGATGAGAGG 1154
Qy 1081 TTTGGGCTTGACACCCAGCATTTGTCTCTGAATGAACTGTGCTCAGTTCAGTTCACAACTT 1140
Db 1155 TTTGGGCTTGACACCCAGCATTTGTCTCTGAATGAACTGTGCTCAGTTCAGTTCACAACTT 1214
Qy 1141 GCGAGCCACAAGATTCAAAACCCAGCTCAACCTCATCTCCACCGGACATCTTCCCTCGTGC 1200
Db 1215 GCGAGCCACAAGATTCAAAACCCAGCTCAACCTCATCTCCACCGGACATCTTCCCTCGTGC 1274
Qy 1201 ACCAGTTTCCGCTGTAAGAGAGGGCCCCACCTCAGTGTGCCATGGTTCAAGGTA 1260
Db 1275 ACCAGTTTCCGCTGTAAGAGAGGGCCCCACCTCAGTGTGCCATGGTTCAAGGTA 1334
Qy 1261 TGCTCTCTCAGTACAGCTCCGCTCCAGAGGGAGTGGCAGAGGATGCCATTATTACT 1320
Db 1335 TGCTCTCTCAGTACAGCTCCGCTCCAGAGGGAGTGGCAGAGGATGCCATTATTACT 1394
Qy 1321 TGCAATCTCTAGGAATTCATAGTTGAGGCGCTGCAAGCTTCCCAACTTCCAGCAGAGCGTG 1380

Db 1395 TGCAATCTCTAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCCAGCAGAGCGTG 1454
Qy 1381 CAGAGTACAGGAGGAGTGGCAGGACGGCCCCAGCCCCAGCAGAGAAAGAGTCAGTAC 1440
Db 1455 CAGAGTACAGGAGGAGTGGCAGGACGGCCCCAGCCCCAGCAGAGAAAGAGTCAGTAC 1514
Qy 1441 CCAGAAATCATCTTCTTGGAAACAGGCTTGCCTATCCGATGGAAGATTGCAAAATGTCACT 1500
Db 1515 CCAGAAATCATCTTCTTGGAAACAGGCTTGCCTATCCGATGGAAGATTGCAAAATGTCACT 1574
Qy 1501 GCCACACTTGTCAAATAAGCCCCGACACAGTCTCTGCTACTTGGACTGTGGTGGAGGACA 1560
Db 1575 GCCACACTTGTCAAATAAGCCCCGACACAGTCTCTGCTACTTGGACTGTGGTGGAGGACA 1634
Qy 1561 TTTGGGACAGTGTGCCGTCAATTACGGAGACAGGTGACAGGGTCTCTGGGACACCTTGCT 1620
Db 1635 TTTGGGACAGTGTGCCGTCAATTACGGAGACAGGTGACAGGGTCTCTGGGACACCTTGCT 1694
Qy 1621 GCTGTGTTGTGTCCTCCACCTGCACGACAGATCACACACGGGCTTGCCTCAAGTATCTTGTCTG 1680
Db 1695 GCTGTGTTGTGTCCTCCACCTGCACGACAGATCACACACGGGCTTGCCTCAAGTATCTTGTCTG 1754
Qy 1681 CAGAGAAACGCGCTTGGCATCTTTGGGAAAGCGCTTCAACCTTTGCTGGTGTGCTGCC 1740
Db 1755 CAGAGAAACGCGCTTGGCATCTTTGGGAAAGCGCTTCAACCTTTGCTGGTGTGCTGCC 1814
Qy 1741 CCCAACAGCTCAAAGCTGGCTCCAGCAGTACCAACAACAGTCCAGGAGGCTCTGCAC 1800
Db 1815 CCCAACAGCTCAAAGCTGGCTCCAGCAGTACCAACAACAGTCCAGGAGGCTCTGCAC 1874
Qy 1801 CACATCAGTATGATCTCTGCCAAATGCTTTCAGGAAGGGCTGAGATCTCCAGTCTCTGCA 1860
Db 1875 CACATCAGTATGATCTCTGCCAAATGCTTTCAGGAAGGGCTGAGATCTCCAGTCTCTGCA 1934
Qy 1861 GTGAAAGATTGATCAGTTCGCTTTCGCAACATGTGATTTTGGAAAGATTTCAGACCTGT 1920
Db 1935 GTGAAAGATTGATCAGTTCGCTTTCGCAACATGTGATTTTGGAAAGATTTCAGACCTGT 1994
Qy 1921 CTGTGGGCACTGCAAGCATGCTTTGGCTGTGGCTGTGTGCACACCTCTGGCTGGAA 1980
Db 1995 CTGTGGGCACTGCAAGCATGCTTTGGCTGTGGCTGTGTGCACACCTCTGGCTGGAA 2054
Qy 1981 GTGTCTTATTTCCGGGACACCATGCCCTGCGAGCTCTGTGTCGGATGGGAAAGATGCC 2040
Db 2055 GTGTCTTATTTCCGGGACACCATGCCCTGCGAGCTCTGTGTCGGATGGGAAAGATGCC 2114
Qy 2041 ACCCTCTGTATACATGAAGCCACCTTGGAAAGATGTTTGGAAAGAGACAGTGGAAAG 2100
Db 2115 ACCCTCTGTATACATGAAGCCACCTTGGAAAGATGTTTGGAAAGAGACAGTGGAAAG 2174
Qy 2101 ACACAGCACACAGTCCCAAGCCATCAGCTGGGGATGCGGATGAACGCGGAGTTCAAT 2160
Db 2175 ACACAGCACACAGTCCCAAGCCATCAGCTGGGGATGCGGATGAACGCGGAGTTCAAT 2234
Qy 2161 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGTCCCTCTTTCAGCCCACTTCACTCAGC 2220
Db 2235 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGTCCCTCTTTCAGCCCACTTCACTCAGC 2294
Qy 2221 GAGAAAGTGGAGTTGGCTTTTGACACATGAAGTCTGCTTTGGAGACTTTTCCAACTATG 2280
Db 2295 GAGAAAGTGGAGTTGGCTTTTGACACATGAAGTCTGCTTTGGAGACTTTTCCAACTATG 2354
Qy 2281 CCCAAGCTGATTTCCCACTGAAAGCCCTGTTGCTGGCGACATCGAGGAGATGGAGAG 2340
Db 2355 CCCAAGCTGATTTCCCACTGAAAGCCCTGTTGCTGGCGACATCGAGGAGATGGAGAG 2414
Qy 2341 GCGAGGAGAAAGCGGAGCTCGCGCAGGTGGCGGGCCCTCTCTGTCCAGGAGCTGCGCA 2400
Db 2415 GCGAGGAGAAAGCGGAGCTCGCGCAGGTGGCGGGCCCTCTCTGTCCAGGAGCTGCGCA 2474
Qy 2401 GCGGCTCTGGAGATGGGGAGCTTCAGCAGAAAGCGGGCCCAACAGAGGAGCCACAGGCC 2460
Db 2475 GCGGCTCTGGAGATGGGGAGCTTCAGCAGAAAGCGGGCCCAACAGAGGAGCCACAGGCC 2534

Db 961 GATCTCTGCTGCTTTTGTGGTGTAGATGTCCAGATGAAGCTTATTCAACCCATC 1020
Qy 961 TGTGAAATGCCACCTTTTCAGAGTACCAAGAAAGGAGATGCCCGCTGGCTTGTG 1020
Db 1021 TGTGAAATGCCACCTTTTCAGAGTACCAAGAAAGGAGATGCCCGCTGGCTTGTG 1080
Qy 1021 GTTTCATGTGGCCCGCAGCATCTGTGCTTGTGGACAGAGGTACCAAGAGTGTGAGAGG 1080
Db 1081 GTTTCATGTGGCCCGCAGCATCTGTGCTTGTGGACAGAGGTACCAAGAGTGTGAGAGG 1140
Qy 1081 TTTTGGCCCTGACACCCAGCAGTCTGTGCTGCTGAATGAGAACTGTGCCCTCAGTTCACAACTT 1140
Db 1141 TTTTGGCCCTGACACCCAGCAGTCTGTGCTGCTGAATGAGAACTGTGCCCTCAGTTCACAACTT 1200
Qy 1141 CGAGCCCAAGATTCAACCCAGCTCAACCTTCATCCACCGGACATTTTCCCCCTGCTC 1200
Db 1201 CGAGCCCAAGATTCAACCCAGCTCAACCTTCATCCACCGGACATTTTCCCCCTGCTC 1260
Qy 1201 ACCAGTTTCCGCTGTAAGAAAGGAGGGCCCAACCTTCAGTGTGCCATGGTTTCAGGGTCAA 1260
Db 1261 ACCAGTTTCCGCTGTAAGAAAGGAGGGCCCAACCTTCAGTGTGCCATGGTTTCAGGGTCAA 1320
Qy 1261 TGCCTCTCAAGTACCAAGTCCGCTCCAGAGGGAGTGGCAGAGGGATGCCATTATTACT 1320
Db 1321 TGCCTCTCAAGTACCAAGTCCGCTCCAGAGGGAGTGGCAGAGGGATGCCATTATTACT 1380
Qy 1321 TGCAATCTGTAGGAATTCATAGTTGAGCGCTGTCAGCTTCCCAATTCAGAGAGCGTG 1380
Db 1381 TGCAATCTGTAGGAATTCATAGTTGAGCGCTGTCAGCTTCCCAATTCAGAGAGCGTG 1440
Qy 1381 CAGGAGTACAGGAGAGTGCAGCAGAGCGGCCAGCCAGCCAGCAGAGAGAAAGTCAAGTAC 1440
Db 1441 CAGGAGTACAGGAGAGTGCAGCAGAGCGGCCAGCCAGCCAGCAGAGAGAAAGTCAAGTAC 1500
Qy 1441 CCAGAAATCATCTTCTTGGAAACAGGGTCTGCCATCCCGATCCGATTCGAAATCTCAGT 1500
Db 1501 CCAGAAATCATCTTCTTGGAAACAGGGTCTGCCATCCCGATTCGAAATCTCAGT 1560
Qy 1501 GCCACATCTGTCAACATTAAGCCCGACACGCTCTCTGTACTTGGATCTGGTGGAGGCACA 1560
Db 1561 GCCACATCTGTCAACATTAAGCCCGACACGCTCTCTGTACTTGGATCTGGTGGAGGCACG 1620
Qy 1561 TTTTGGCAGCTGTGCGCTCATTTACGAGACCAAGTGGACAGGGTCTCGGCACCCCTGGCT 1620
Db 1621 TTTTGGCAGCTGTGCGCTCATTTACGAGACCAAGTGGACAGGGTCTCGGCACCCCTGGCT 1680
Qy 1621 GCTGTGTTTGTGTCCTCCACTGCAACGAGATCACCAACGGGCTTCCCAAGTATCTTGTG 1680
Db 1681 GCTGTGTTTGTGTCCTCCACTGCAACGAGATCACCAACGGGCTTCCCAAGTATCTTGTG 1740
Qy 1681 CAGAGAGAACCGCCTTGGCATCTTTTGGAAAGCGCTTACCCCTTGTGCTGGTGTGCC 1740
Db 1741 CAGAGAGAACCGCCTTGGCATCTTTTGGAAAGCGCTTACCCCTTGTGCTGGTGTGCC 1800
Qy 1741 CCCAACGAGCTCAAGCCTGGCTCCAGCAGTACCAACAGTGCAGAGGCTCTGTCAC 1800
Db 1801 CCCAACGAGCTCAAGCCTGGCTCCAGCAGTACCAACAGTGCAGAGGCTCTGTCAC 1860
Qy 1801 CACATCAGTATGATTCCTGCCAAATTCCTTACGAGAGGGCTGAGATCTCCAGTCTGCA 1860
Db 1861 CACATCAGTATGATTCCTGCCAAATTCCTTACGAGAGGGCTGAGATCTCCAGTCTGCA 1920
Qy 1861 GTGGAAGATTGATCAGTTCGCTGTGGCAACATGTGATTTGGAAGAGTTTCAGACCTGT 1920
Db 1921 GTGGAAGATTGATCAGTTCGCTGTGGCAACATGTGATTTGGAAGAGTTTCAGACCTGT 1980
Qy 1921 CTGGTGGCAGTCAAGCATCGTTTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGAA 1980
Db 1981 CTGGTGGCAGTCAAGCATCGTTTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGAA 2040
Qy 1981 GTGGTCTATTCCGGGGACACCATGCTTCCGAGGCTCTGGTCCGGATGGGAAAGATGCC 2040

Db 2041 GTGCTATTCCGGGGACACCATGCCCTCGAGGCTCTGCTCGGATGGGAAAGATGCC 2100
Qy 2041 ACCCTCTGATACATGAAGCCACCCTCGAAGATGGTTTGGNAGGAGAGTGGAAAG 2100
Db 2101 ACCCTCTGATACATGAAGCCACCCTCGAAGATGGTTTGGNAGGAGAGTGGAAAG 2160
Qy 2101 ACACACAGCAACGCTCCCAAGCCATCAGCGTGGGGATGCGGATGAACGCGGAGTTCAAT 2160
Db 2161 ACACACAGCAACGCTCCCAAGCCATCAGCGTGGGGATGCGGATGAACGCGGAGTTCAAT 2220
Qy 2161 ATGCTGAACCACTTTCAGCCAGCGCTATGCCAAGTCCCGCTTTCAGCCCACTTCAAG 2220
Db 2221 ATGCTGAACCACTTTCAGCCAGCGCTATGCCAAGTCCCGCTTTCAGCCCACTTCAAG 2280
Qy 2221 GAGAAAGTGGAGTTGCTTTTGACCATGAAGTCTGCTTTGGAGACTTTTCCAACTATG 2280
Db 2281 GAGAAAGTGGAGTTGCTTTTGACCATGAAGTCTGCTTTGGAGACTTTTCCAACTATG 2340
Qy 2281 CCCAAGCTGATTTCCCACTGAAAGCCCTGTTTCTCGGCACATCGAGGAGATGGAGG 2340
Db 2341 CCCAAGCTGATTTCCCACTGAAAGCCCTGTTTCTCGGCACATCGAGGAGATGGAGG 2400
Qy 2341 CGCAGGAGAAAGCGGGAGTGCAGGAGTGCAGGAGTGCAGGAGTGCAGGAGTGCAGG 2400
Db 2401 CGCAGGAGAAAGCGGGAGTGCAGGAGTGCAGGAGTGCAGGAGTGCAGGAGTGCAGG 2460
Qy 2401 GCGGCTGTGAGGATGGGAGCCCTCAGCAGAGAGGGGGCCACACAGAGGAGCCAGGCC 2460
Db 2461 GCGGCTGTGAGGATGGGAGCCCTCAGCAGAGAGGGGGCCACACAGAGGAGCCAGGCC 2520
Qy 2461 AAGAAGTCTCAGAGCCCACTGTA 2481
Db 2521 AAGAAGTCTCAGAGCCCACTGTA 2541

BC004158 3006 bp mRNA linear PRI 12-JUL-2001
Homo sapiens, putative prostate cancer susceptibility protein,
clone MGC:2441 IMAGE:2820640, mRNA, complete cds.
BC004158
ACCESSION BC004158.1 GI:13278770
VERSION BC004158.1
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 3006)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapsb@mail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
<http://www.systemsbio.org>
contact: amadan@systemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia Greene, Mark Kettman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 2 Row: C Column: 13
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7022621.
Location/Qualifiers
1. .3006

Qy 1681 CAGAGAGACGGCCCTTGGCATCTTTGGGAAGCGCTTCAACCTTTGCTGGTGTGCC 1740
Db 1750 CAGAGAGACGGCCCTTGGCATCTTTGGGAAGCGCTTCAACCTTTGCTGGTGTGCC 1809
Qy 1741 CCCAACCAAGCTCAAAAGCTGGCTCCAGCAGTACCAACACCAAGTGCAGAGGCTCGCAC 1800
Db 1810 CCCAACCAAGCTCAAAAGCTGGCTCCAGCAGTACCAACACCAAGTGCAGAGGCTCGCAC 1869
Qy 1801 CACATCAGTATGATTCCTGCCAAATGCCTTCAGGAAGGGCTGAGATCTCCAGTCTCGCA 1860
Db 1870 CACATCAGTATGATTCCTGCCAAATGCCTTCAGGAAGGGCTGAGATCTCCAGTCTCGCA 1929
Qy 1861 GTGGAAGATGATCAGTTCGCTGTGGCAATGATGATTTGGAGAGTTCAGACCTGT 1920
Db 1930 GTGGAAGATGATCAGTTCGCTGTGGCAATGATGATTTGGAGAGTTCAGACCTGT 1989
Qy 1921 CTGGTGGCGCACTCAAGCATGCTTTGGCTGTGGCTGGTGCACACCTCTGGCTGGAAA 1980
Db 1990 CTGGTGGCGCACTCAAGCATGCTTTGGCTGTGGCTGGTGCACACCTCTGGCTGGAAA 2049
Qy 1981 GTGGTCTATTCCGGGACACCATGCTTCGAGGCTCTGGTCCGATGGGAAAGATGCC 2040
Db 2050 GTGGTCTATTCCGGGACACCATGCTTCGAGGCTCTGGTCCGATGGGAAAGATGCC 2109
Qy 2041 ACCCTCTGATACATGAAGCACCCTGGAAGATGTTTGGAGAGGAGCATGGAAAG 2100
Db 2110 ACCCTCTGATACATGAAGCACCCTGGAAGATGTTTGGAGAGGAGCATGGAAAG 2169
Qy 2101 ACACAGACACAACTCCCAAGCCATCAGCGTGGGATGCGGATCAACCGGAGTTCATT 2160
Db 2170 ACACAGACACAACTCCCAAGCCATCAGCGTGGGATGCGGATCAACCGGAGTTCATT 2229
Qy 2161 ATGCTGAACCACTTCAGCAGCGCTATGCCAAGTTCCTCTTCAGCCCAACTTCAGC 2220
Db 2230 ATGCTGAACCACTTCAGCAGCGCTATGCCAAGTTCCTCTTCAGCCCAACTTCAGC 2289
Qy 2221 GAGAAAGTGGGAGTTCCTTTGACACATGAGTCTGCTTTGGAGATTTCCACAAATG 2280
Db 2290 GAGAAAGTGGGAGTTCCTTTGACACATGAGTCTGCTTTGGAGATTTCCACAAATG 2349
Qy 2281 CCCAAGCTGATTCCTCCCACTGAAAGCCCTGTTGCTGGCGACATCGAGGATGAGGAG 2340
Db 2350 CCCAAGCTGATTCCTCCCACTGAAAGCCCTGTTGCTGGCGACATCGAGGATGAGGAG 2409
Qy 2341 CCGAGGAGAAAGCGGAGCTGCGCAGTGGCGGCGGCTCTCTTCAGGAGCTGCA 2400
Db 2410 CCGAGGAGAAAGCGGAGCTGCGCAGTGGCGGCGGCTCTCTTCAGGAGCTGCA 2469
Qy 2401 GCGGCTGGAGGATGGGAGCTTCAGCAGAGCGGCGCCACACAGAGAGCCACAGGCC 2460
Db 2470 GCGGCTGGAGGATGGGAGCTTCAGCAGAGCGGCGCCACACAGAGAGCCACAGGCC 2529
Qy 2461 AAGAAAGTCAGAGCCAGTGA 2481
Db 2530 AAGAAAGTCAGAGCCAGTGA 2550

RESULT 5
AF308698
LOCUS AF308698 2908 bp mRNA linear PRI 27-FEB-2001
DEFINITION Pan troglodytes ELAC2 mRNA, complete cds.
ACCESSION AF308698
VERSION AF308698.1 GI:10946496
KEYWORDS
SOURCE Pan troglodytes.
ORGANISM Pan troglodytes.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
REFERENCE
1 (bases 1 to 2908)
Tavtigian,S.V., Simard,J., Teng,D.H.F., Abtin,V., Baumgard,M.,
Beck,A., Camp,N.J., Carillo,A.R., Chen,Y., Dayananth,P.,
Desrochers,M., Dumont,M., Farnham,J.M., Frank,D., Frye,C.,
Ghaffari,S., Gupte,J.S., Hu,R., Iliev,D., Janacki,T., Kort,E.N.,
Leaity,K., Leavitt,A., Leblanc,G., McArthur-Morrison,J.,
Pederson,A., Penn,B., Peterson,K.T., Reid,J.E., Richards,S.,
Schroeder,M., Smith,R., Snyder,S.C., Swedlund,B., Swensen,J.,
Thomas,A., Tranchant,M., Woodland,A.-M., Labrie,F., Skolnick,M.H.,
Neuhausen,S., Rommens,J. and Cannon-Albright,L.A.
A candidate prostate cancer susceptibility gene at chromosome 17p
Nat. Genet. 27 (2), 172-180 (2001)
21096977
MEDLINE
PUBMED
AUTHORS

Ghaffari,S., Gupte,J.S., Hu,R., Iliev,D., Janacki,T., Kort,E.N.,
Leaity,K., Leavitt,A., Leblanc,G., McArthur-Morrison,J.,
Pederson,A., Penn,B., Peterson,K.T., Reid,J.E., Richards,S.,
Schroeder,M., Smith,R., Snyder,S.C., Swedlund,B., Swensen,J.,
Thomas,A., Tranchant,M., Woodland,A.-M., Labrie,F., Skolnick,M.H.,
Neuhausen,S., Rommens,J. and Cannon-Albright,L.A.
A candidate prostate cancer susceptibility gene at chromosome 17p
Nat. Genet. 27 (2), 172-180 (2001)
21096977
MEDLINE
PUBMED
AUTHORS

TITLE
JOURNAL
REFERENCE

TITLE
JOURNAL
REFERENCE

TITLE
JOURNAL
REFERENCE

FEATURES
source
1..2908
/organism="Pan troglodytes"
/db xref="taxon:9598"
1..2481
/codon_start=1
/product="ELAC2"
/protein_id="AAG24920.1"
/db xref="GI:10946497"
/translation="MWALCSLLRSAAAGTMSQGRITISOAPRRERPKDPLRLHRTRE
KPGSCSGGPNVLYQVVAAGSRDSGAALYVSEFNRYLFCNGEGIQRLMRHKLKV
ARLDNLFELRMHNSVGGSLNILLTKETGLPKCVLGGPQLEKYLEAIKIFSGPLKG
IELAUPHSAPEYEDETIVYQIPHSQRRGKHQWQSPERPLSRSPERSDSSEN
ENEPHLPHGVQSRGVRDSSLVAFIKLHKGKGNFLVAKEMGLPVGTAATPIIA
AVKDKSITHEGREILABELTTPDPGAAFPVVECPDESGFIQICENATFORQKAD
APVALVHMVAPESLVDSRVQMMERFGPDTHLVNENCASVHNLRSKHQIOTQLNI
HPDIFLLTSFPCKKEGPTLSVPMVOGECLLKYQLPRRERORDAITCNPEFLIEA
LQPNFQQSVQETRRSAQGPAPAEKRSQYPEILFLTGSAIPMKIRNVSATLVNISF
DTSLLDCGEGTGGQLCRHYGDQVDRVLGTLAAVFVSHLHHTGLNLLQRRAL
ASLGRPFHLLVAVPNQLKAWQYHNOQOEVLHHSMTIAGKLOEAGELSSPAVERL
ISSLRATCDLEEFOTCLVRCHKAFGCAVHTSKWVYSGDTPMCEALVRMGKDATL
LIHEATLEDGLBEEAVEKTHSTTSQAI SVGMNNAEFLIMLHFSQRYAKVPLFSPNF
LKGVLAPDHMKVCFGDFATMPKLIPLKALFAGDIEMEERERREKRELQVRAALLSRE
LKGVLGDEGPPQKRATPEEPQAKVRAQ"

BASE COUNT 712 a 788 c 819 g 589 t
ORIGIN

Query Match 99.0%; Score 2455.4; DB 9; Length 2908;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 2465; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 1 ATGTGGGCGCTTTGCTCGCTGCTCGGTCGCGCGCGGACGACCATGTTCGACGAGCGC 60
Db 1 ATGTGGGCGCTTTGCTCGCTGCTCGGTCGCGCGCGGACGACCATGTTCGACGAGCGC 60
Qy 61 ACCATATCGAGGACCGC 120
Db 61 ACCATATCGAGGACCGC 120
Qy 121 CGCAGCGGAGAGAGCGCGGACCGCTCGGGGTGCTCGCGCGCGCGCGCGCGCGCGCG 180

Db 121 CCGCAGGAGAGCGGGAGCCGTCGGGGTGCTCCGGCGGCCCAACACCGGTGACCTG 180
Qy 181 CAGGTGGTGGCAGCGGGTAGCCGGAGCTCGGGCGCGGGCTCTAGCTCTTCTCCAGTTTC 240
Db 181 CAGGTGGTGGCAGCGGGTAGCCGGAGCTCGGGCGCGGGCTCTAGCTCTTCTCCAGTTTC 240
Qy 241 AACCGGTATCTCTTCAACTGTGGAGAGCGGTTTCAGAGACTCATGAGAGACAAAGTTA 300
Db 241 AACCGGTATCTCTTCAACTGTGGAGAGCAATTCAGAGACTCATGAGAGACAAAGTTA 300
Qy 301 AAGGTGGTGGTGGCAGCAACATATTCCTGACACGAATGCACTGCTCTAAATGTTGGGGC 360
Db 301 AAGGTGGTGGTGGCAGCAACATATTCCTGACACGAATGCACTGCTCTAAATGTTGGGGC 360
Qy 361 TTAAGTGGAAATGATTTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTAATTTCTGGA 420
Db 361 TTAAGTGGAAATGATTTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTAATTTCTGGA 420
Qy 421 CTTCCAACTGGGAAATATCTCTGAGCAATCAAAATATTTTCTGGTCCATTTGAAAGGA 480
Db 421 CTTCCAACTGGGAAATATCTCTGAGCAATCAAAATATTTTCTGGTCCATTTGAAAGGA 480
Qy 481 ATAGAACTGGCTGGCGGCCCACTCTGCCAGAAATACGAGATGCAACATGACAGTT 540
Db 481 ATAGAACTGGCTGGCGGCCCACTCTGCCAGAAATACGAGATGCAACATGACAGTT 540
Qy 541 TACCAGATCCCATACACAGTGAACAGAGAGGGGAAAGCACCAACATGGCAGAGTCCA 600
Db 541 TACCAGATCCCATACACAGTGAACAGAGAGGGGAAAGCACCAACATGGCAGAGTCCA 600
Qy 601 GAAAGCCCTCTCAGCAGGCTCAGTCCAGAGCGATCTTACAGTCCGAGTCAAAAT 660
Db 601 GAAAGCCCTCTCAGCAGGCTCAGTCCAGAGCGATCTTACAGTCCGAGTCAAAAT 660
Qy 661 GAGCACACCTTCCATGGTGTAGCCAGAGAGAGGGGTGAGGACTCTTCCCTGGTC 720
Db 661 GAGCACACCTTCCATGGTGTAGCCAGAGAGAGGGGTGAGGACTCTTCCCTGGTC 720
Qy 721 GTAGCTTTTCACTGTAAAGTTCACCTTAAAGAGAGAACTCTTGGTGTCAAAGCAAG 780
Db 721 GTAGCTTTTCACTGTAAAGTTCACCTTAAAGAGAGAACTCTTGGTGTCAAAGCAAG 780
Qy 781 GAGATGGGCTCCAGTGGGACAGCTGCCATCGCTCCCATCTGCTGTCAAGGAC 840
Db 781 GAGATGGGCTCCAGTGGGACAGCTGCCATCGCTCCCATCTGCTGTCAAGGAC 840
Qy 841 GGGAAAGCATCACTCATGAAGAGAGAGATTTGGCTGAAGAGCTGTACTCTCCA 900
Db 841 GGGAAAGCATCACTCATGAAGAGAGAGATTTGGCTGAAGAGCTGTACTCTCCA 900
Qy 901 GATCTGTGTGCTTTTGTGGTGTAGAAATGTCCAGATGAAAGCTTCATTCAACCCATC 960
Db 901 GATCTGTGTGCTTTTGTGGTGTAGAAATGTCCAGATGAAAGCTTCATTCAACCCATC 960
Qy 961 TGTGAAATGCCACCTTTTCCAGAGTACCAAGGAAAGGAGATGCCCGGTGGCTTGGTG 1020
Db 961 TGTGAAATGCCACCTTTTCCAGAGTACCAAGGAAAGGAGATGCCCGGTGGCTTGGTG 1020
Qy 1021 GTTCATGGCCCGCAGCATCTGTGCTTGTGGACAGCAGGTATCAGCAGTGGATGGAGG 1080
Db 1021 GTTCATGGCCCGCAGAAATCTGTGCTTGTGGACAGCAGGTATCAGCAGTGGATGGAGG 1080
Qy 1081 TTTGGSCCTGACACCCAGCACTTGTCTGTGAATGAGAACTGTGCTCAGTTCACAACTT 1140
Db 1081 TTTGGSCCTGACACCCAGCACTTGTCTGTGAATGAGAACTGTGCTCAGTTCACAACTT 1140
Qy 1141 CGCAGCCAAAGATTCAAAACCCAGCTCAACCTCATCCACCGGACATCTTCCCTGTGTC 1200
Db 1141 CGCAGCCAAAGATTCAAAACCCAGCTCAACCTCATCCACCGGACATCTTCCCTGTGTC 1200
Qy 1201 ACCAGTTTCCGTGTAAGAGAGGGGCCCAACCTCAGTGTGCCATGTTTTCAGGGTGA 1260

Db 1201 ACCAGTTTCCCTGTAAAGAGGGGCCCAACCTCAGTGTGCCATGTTTCAGGGTGA 1260
Qy 1261 TGCCTCTCTCAAGTACCAGTCCGTCCAGAGGAGGTGGCAGAGGATGCCATTATTACT 1320
Db 1261 TGCCTCTCTCAAGTACCAGTCCGTCCAGAGGAGGTGGCAGAGGATGCCATTATTACT 1320
Qy 1321 TGCATCTCTGAGGAATTCATAGTTGAGGCGCTGAGCTTCCCACTTCCAGCAGAGCGTG 1380
Db 1321 TGCATCTCTGAGGAATTCATAGTTGAGGCGCTGAGCTTCCCACTTCCAGCAGAGTG 1380
Qy 1381 CAGGAGTACAGAGGAGTGGCAGGACGCGCCAGCCCGCAGCAGAGAGAAAGTCAGTAC 1440
Db 1381 CAGGAGTACAGAGGAGTGGCAGGACGCGCCAGCCCGCAGCAGAGAGAAAGTCAGTAC 1440
Qy 1441 CCAGAAATCATCTTCTTTGGAACAGGGTCTGCCATCCCGATGAAGATTGAAATGTCA 1500
Db 1441 CCAGAAATCATCTTCTTTGGAACAGGGTCTGCCATCCCGATGAAGATTGAAATGTCA 1500
Qy 1501 GCCACACTTGTCAACATGAAGCCCGCAGACGCTCTCTGCTACTGGACTGTGGTGGGCA 1560
Db 1501 GCCACACTTGTCAACATGAAGCCCGCAGACGCTCTCTGCTACTGGACTGTGGTGGGCA 1560
Qy 1561 TTTGGGCACTGTGCGGTCATTACGGAGACCAAGGTGGACAGGGTCTCGGCAACCTGGCT 1620
Db 1561 TTTGGGCACTGTGCGGTCATTACGGAGACCAAGGTGGACAGGGTCTCGGCAACCTGGCT 1620
Qy 1621 GCTGTGTTTGTGTCCCACTGCACGAGATCACACAGGGCTTCCCAAGTATCTTGTG 1680
Db 1621 GCTGTGTTTGTGTCCCACTGCACGAGATCACACAGGGCTTCCCAAGTATCTTGTG 1680
Qy 1681 CAGAGAGAACGGCTTGGCATCTTTGGGAAAGCGCTTACCCCTTGTGGTGGTGGCT 1740
Db 1681 CAGAGAGAACGGCTTGGCATCTTTGGGAAAGCGCTTACCCCTTGTGGTGGTGGCT 1740
Qy 1741 CCCAACAGCTCAAGCGCTGCCAGCAGTACCAACACAGTGCAGAGGTCTCTGCAC 1800
Db 1741 CCCAACAGCTCAAGCGCTGCCAGCAGTACCAACACAGTGCAGAGGTCTCTGCAC 1800
Qy 1801 CACATCAGTATGATTCTCTCCAAATGCCTTCAGGAAGGGGTGAGATCTCCAGTCTGCA 1860
Db 1801 CACATCAGTATGATTCTCTCCAAATGCCTTCAGGAAGGGGTGAGATCTCCAGTCTGCA 1860
Qy 1861 GTGGAAGATGATCAGTTCCGTGTGGCAACATGTGATTTGGAAAGATTTTCAGACCTGT 1920
Db 1861 GTGGAAGATGATCAGTTCCGTGTGGCAACATGTGATTTGGAAAGATTTTCAGACCTGT 1920
Qy 1921 CTGTTGCGGCACTGCAAGCATGCGTTTGGCTGTGCGTGGTGCACACCTCTCGCTGGA 1980
Db 1921 CTGTTGCGGCACTGCAAGCATGCGTTTGGCTGTGCGTGGTGCACACCTCTCGCTGGA 1980
Qy 1981 GTGTTCTATTCCGGGACACCATGCGCTCGAGGCTCTGGTCCGGATGGGAAAGTGC 2040
Db 1981 GTGTTCTATTCCGGGACACCATGCGCTCGAGGCTCTGGTCCGGATGGGAAAGTGC 2040
Qy 2041 ACCCTCTGATACATGAAGCCACCTTGGAAAGCGTTTGGAAAGAGAGTGGAAAG 2100
Db 2041 ACCCTCTGATACATGAAGCCACCTTGGAAAGCGTTTGGAAAGAGAGTGGAAAG 2100
Qy 2101 ACACAGCACAACGTCCTCCAGCCATCAGCGTGGGATGCGGATCAACCGGAGTTCATT 2160
Db 2101 ACACAGCACAACGTCCTCCAGCCATCAGCGTGGGATGCGGATCAACCGGAGTTCATT 2160
Qy 2161 ATGCTGAAACCACTTCAGCCAGCGCTATGCCAAGTCCCTCTTTCAGCCCCCACTTCA 2220
Db 2161 ATGCTGAAACCACTTCAGCCAGCGCTATGCCAAGTCCCTCTTTCAGCCCCCACTTCA 2220
Qy 2221 GAGAAAGTGGAGTTCCTTTGACCAATGAAGTCTGCTTTGGAGATCTTTCACCAATG 2280
Db 2221 GAGAAAGTGGAGTTCCTTTGACCAATGAAGTCTGCTTTGGAGATCTTTCACCAATG 2280
Qy 2281 CCCAAGCTGATTCCTCCCACTGAAAGCCCTGTTTGTGGGACATCGAGAGATGGAGG 2340
Db 2281 CCCAAGCTGATTCCTCCCACTGAAAGCCCTGTTTGTGGGACATCGAGAGATGGAGG 2340

```
Qy 2341 GCAGGAGAGACGGAGCTCGGCGAGTGGCGGGCCCTCTGTCAGGAGCTGGCA 2400
Db 2341 GCAGGAGAGACGGAGCTCGGCGAGTGGCGGGCCCTCTGTCAGGAGCTGGCA 2400
Qy 2401 GCGGCGCTGGAGGATGGGAGCTCAGCAGAGAGCGGGCCACACAGAGAGAGCCAGAGGCC 2460
Db 2401 GCGGCGCTGGAGGATGGGAGCTCAGCAGAGAGCGGGCCACACAGAGAGAGCCAGAGGCC 2460
Qy 2461 AAGAGGTCAGAGCCAGTGA 2481
Db 2461 AAGAGGTCAGAGCCAGTGA 2481

RESULT 6
AK001392 Homo sapiens cDNA FLJ10530 fis, clone NT2RP2000985. linear PRI 01-AUG-2002
LOCUS AK001392 2976 bp mRNA
DEFINITION Homo sapiens cDNA FLJ10530 fis, clone NT2RP2000985.
ACCESSION AK001392
VERSION AK001392.1 GI:7022621
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens teratocarcinoma cell line:NT2 cDNA to mRNA,
clone lib:NT2RP2 clone:NT2RP2000985.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
Wagatsuma,M., Hosoi,T., Kaku,Y., Kido,H., Kondo,H.,
Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,
Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,
Nakamura,Y., Nagahara,K., Masuo,Y., Ninomiya,K. and Iwayanagi,T.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2976)
AUTHORS Isogai,T. and Otsuki,T.
JOURNAL Direct Submission
SUBMITTED (16-FEB-2000) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5'- and 3'-end one pass sequencing and clone selection:
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
source Location/Qualifiers
1..2976
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2RP2000985"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_libs="NT2RP2"
/feature="cloning vector: pME18SFL3-mRNA from NT2 neuronal
precursor cells after 2-weeks retinoic acid (RA)
induction."
75..2639
/feature="unnamed protein product"
/codon_start=1
/protein_id="BAA91666.1"
/db_xref="GI:7022622"
/translation="MWALCSLLRLAAGRTMSOGRITISQAPARRPRPKDLRLHRTRE
KRPGSGGNTVYLVVAVAGSDSGAALVVFSEFNRYLPNCEGVQRLMOEHKLAV
ARLDNIPLTRMHWNVGGLSMILTKETGLPKVLSGPOLKYLKAIKFSPLKX
TELAWRHPSEYDEFTVYQIFHSEQRGRQWPQSPERLUSRPSKSSDSN
ENEPHLPHGVSRQGRVDSLSLVAFIKLHLKRNFLVAKEMGLPVGTAATAPIIA
AVDKGKSTHREGRLAEELCTPPDGAFAVVECPDESEFTQPCENATFORVQKAD
APVALVVMAPASVLDVSRVQWMERFGPDQHLVNLNENCASVHNLRSKHTQTLNLI
HPDIFPLLTSPRCKEGPTLSVPMVQGECLLKYLQLRPRWQRDAIITCNPEEFVEAL
```

```
QLPNFOOSVQVEYRRSADGPAPAEKRSQYPEIIFLGTSAIPMKIRNVATLVNISPD
TSLLLCGSEGTGOLCRHYGQDOVRLVLTAAVFSVHLHADHTGLPSIILORERALA
SLGKPLRPLLVAPNCLRWLOQYVNCQVEVLHHSIMIPAKCLQEGNAISSPAVERLI
SSLURTDLLEFPOTCLVRICKHAFYGCALVHSGKVVYSGDTWPCALVRMGKDATLL
IHEARTLEDGLBEEAVEKTHSTTSQAI SVGRMNAEFIMLNHFQRYAKVFLFSPNFE
KUGVAFDHMKVCLGDFPTMPKLIPLKALFAGDIEMEEERREKRELRQRAALLSREL
AGGLEDSLSRSQPTQSRPRSEPSSEDLGPELRRLCLVFCPTHAPVSALLAGRS"
BASE COUNT 712 a 807 c 856 g 501 t
ORIGIN
Query Match 98.7%; Score 2448.2; DB 9; Length 2976;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2474; Conservative 0; Mismatches 3; Indels 4; Gaps 2;
Qy 1 ATGTGGGGGCTTTGCTCGCTGCTCGGCTCGCGCGCGGACGACCATGTGCGCAGGAGCC 60
Db 75 ATGTGGGGGCTTTGCTCGCTGCTCGGCTCGCGCGCGGACGACCATGTGCGCAGGAGCC 134
Qy 61 ACCATATCGCAGGACCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
Db 135 ACCATATCGCAGGACCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 194
Qy 121 CGCACGCGAGAGAGCGCGGACCGTCCGGGGTGTCTCCGGCGGCGGCGGCGGCGGCGGCGG 180
Db 195 CGCACGCGAGAGAGCGCGGACCGTCCGGGGTGTCTCCGGCGGCGGCGGCGGCGGCGGCGG 254
Qy 181 CAGTGTGTGCGAGGGGTAGCGGGACTCGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
Db 255 CAGTGTGTGCGAGGGGTAGCGGGACTCGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 314
Qy 241 AACCGGTATCTCTTCAACTGTGGAGAGGGCGTTCAGAGAGATCATGCGAGGAGCACAAAGTTA 300
Db 315 AACCGGTATCTCTTCAACTGTGGAGAGGGCGTTCAGAGAGATCATGCGAGGAGCACAAAGTTA 374
Qy 301 AAGTTGCTCGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
Db 375 AAGTTGCTCGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 434
Qy 361 TTAAGTGAATGATCTTACTTTTAAAGAAACCGGGCTTCCAAAGTGTGTACTTCTTCTGGA 420
Db 435 TTAAGTGAATGATCTTACTTTTAAAGAAACCGGGCTTCCAAAGTGTGTACTTCTTCTGGA 494
Qy 421 CCTCCACAACCTGGAATAATACCTCGAAGCAATCAAAATATTTCTGCTCCATTTGAAAGGA 480
Db 495 CCTCCACAACCTGGAATAATACCTCGAAGCAATCAAAATATTTCTGCTCCATTTGAAAGGA 554
Qy 481 ATAGAACTGGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
Db 555 ATAGAACTGGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 614
Qy 541 TACCAGATCCCATACACAGTGAACAGAGAGGGGAAAGCAACCAACCATGCGCAGAGTCCA 600
Db 615 TACCAGATCCCATACACAGTGAACAGAGAGGGGAAAGCAACCAACCATGCGCAGAGTCCA 674
Qy 601 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCCGAATGAAAT 660
Db 675 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCCGAATGAAAT 734
Qy 661 GAGCCACACCTTCCACATGGTGTGTAGCCAGAGAGAGGGGTGAGGAGCTCTTCCCTGGTC 720
Db 735 GAGCCACACCTTCCACATGGTGTGTAGCCAGAGAGAGGGGTGAGGAGCTCTTCCCTGGTC 794
Qy 721 GTAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGAACTTCTTGGTGTCTCAAGCAAG 780
Db 795 GTAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGAGAACTTCTTGGTGTCTCAAGCAAG 854
Qy 781 GAGATGGGCTCCAGTGTGGGACAGCTGCCATCGCTCCCATCATTTGCTGTCTCAAGGAC 840
Db 855 GAGATGGGCTCCAGTGTGGGACAGCTGCCATCGCTCCCATCATTTGCTGTCTCAAGGAC 914
Qy 841 GGGAAAAGCATCACTCATGAAGAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTCTCCA 900
Db 915 GGGAAAAGCATCACTCATGAAGAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTCTCCA 974
```


JOURNAL Unpublished
REFERENCE 3 (bases 1 to 2893)
AUTHORS Tavtigian,S.V., Simard,J., Teng,D.H.F., Baumgard,M., Beck,A., Camp,N.J., Carillo,A.R., Chen,Y., Dayananth,P., Destrochers,M., Dumont,M., Farnham,J.M., Frank,D., Frye,C., Ghaffari,S., Gupte,J.S., Hu,R., Iliev,D., Janeski,T., Kort,E.N., Laity,K.E., Leavitt,A., Leblanc,G., McArthur-Morrison,J., Pederson,A., Penn,B., Peterson,K.T., Reid,J.B., Richards,S., Schroeder,M., Smith,R., Snyder,S.C., Swedlund,B., Swensen,J., Thomas,A., Tranchant,M., Woodland,A.-M., Labrie,F., Skolnick,M.H., Neuhausen,S., Rommens,J. and Cannon-Albright,L.A.

TITLE Direct Submission
JOURNAL Submitted (27-SEP-2000) Myriad Genetics, Inc., 320 Wakara Way, Salt Lake City, UT 84103, USA

FEATURES
source
1..2893
/organism="Gorilla gorilla"
/db_xref="taxon:9593"
1..2893
/gene="ELAC2"
/gene="ELAC2"
1..2481
/note="similar to Homo sapiens HPC2/ELAC2"
/codon_start=1
/product="ELAC2"
/protein_id="AAG24916.1"
/db_xref="GI:10946489"
/translation="MWALCSLLRSAAQTMSQGRITISQAPARRRPRKDLRLHRLTRE KRPGSGSGPNTVYLQVAAAGSDSAALYVSEFNRYLFCNCEGVORLMOBHLKV VRLDNIILTRMHNVGGLSMILTKETGLPKVGLPGLPPQLEKYLRAIKIPSGPLKG IELAVRPHSAPEYDEFTWYQIPIHSEQRGRHQWQSPERPLSLRSPERSDSEN ENPDLPHGVQRGSDSLVAFI CKLHLKRNFLVKAKEMGLPVTAALAPIIA AVDKKRTIHGREILABELCTPPDGAALFVVECPDESFQIPICENATFORVGKAD APVALVHMAPESVLVPSRYQWMERFQPTQHLVNLNENCAVSHNRSHKIQTLNLI HPDIPLLTGPPSPCKKEPTLSVPMVQGECLLKYLRRPREWORDAIITCNPEFIVEA LQLPNFQSVQVEXRSDVPAPAEKRSQYPEIIFLGTSAPIMKINVSATLVNLSPL DTSLLDCCGEGTFQOLCRHYGDQVDRVLGTLAAVFSVHLRHADHTGLNLLQREQAL ASUGKPLPLVAPSQQLWQVLIHHSIMIPAKLOQGEAEISSFAVERL LISLRLTDLDEEFATCLVRHCKHAFGALVHTSGWKVYVSGDTPMPCBALVRMGKDATL LIHEATLEDGLIEEAVKEHTSTTSQAI SVGNRMNAEFLINLHFSQRYAKVLPFSNPN EKVGVAFDHMKVCFDPTWPKLIPPLKALFAGDIEBMEERREKRELVRQAALLSGE LAGLEDEGEPOQKRAHTEEPQAKVRAQ"

BASE COUNT 705 a 787 c 815 g 586 t

ORIGIN
Query Match 98.5%; Score 2442.6; DB 9; Length 2893;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 2457; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
Qy 1 ATGTGGGCGCTTGTCTGCTGCTGGGTCCGGCGCGGACCGACCATGTCCGAGGAGCG 60
Db 1 ATGTGGGCGCTTGTCTGCTGCTGGGTCCGGCGCGGACCGACCATGTCCGAGGAGCG 60
Qy 61 ACCATATCGAGGACCGCGCGCGGACCGCGCGGACCGCGCGCGCGCGCGCGCGCGCG 120
Db 61 ACCATATCGAGGACCGCGCGCGGACCGCGCGGACCGCGCGGACCGCGCGCGCGCG 120
Qy 121 CGCAGCGAGAGAACCGGACCGTCCGGGTGCTCCGGCGCGCGCGCGCGCGCGCGCGCG 180
Db 121 CGCAGCGAGAGAACCGGACCGTCCGGGTGCTCCGGCGCGCGCGCGCGCGCGCGCG 180
Qy 181 CAGGTGGTGGACGGGTAGCGGACCTCGGGCGCGCGCGCTCTAGCTTCTCCAGATTC 240
Db 181 CAGGTGGTGGACGGGTAGCGGACCTCGGGCGCGCGCGCTCTAGCTTCTCCAGATTC 240
Qy 241 AACCGGTATCTCTTCAACTGTGGAGAGCGTTTACAGACTCATCGACGAGCACAAAGTTA 300
Db 241 AACCGGTATCTCTTCAACTGTGGAGAGCGTTTACAGACTCATCGACGAGCACAAAGTTA 300
Qy 301 AAGGTGTGCTCGCTGGACAACATATTCCTGACACGAATGCACTGGTCTAATGTTGGGGG 360
Db 301 AAGGTGTGCTCGCTGGACAACATATTCCTGACACGAATGCACTGGTCTAATGTTGGGGG 360

Qy 361 TTAAGTGAATGATTCTTACTTTAAAGGAACCGGGCTTCCAAAGTGTGTACTTTCTTGG 420
Db 361 TTAAGTGAATGATTCTTACTTTAAAGGAACCGGGCTTCCAAAGTGTGTACTTTCTTGG 420
Qy 421 CCTCCACAACTGGAAAAATAACCTCGAAGCAATCAAAAATATTTTCTGGTCCATTGAAAGGA 480
Db 421 CCTCCACAGCTGGAAAAATAACCTCGAAGCAATCAAAAATATTTTCTGGTCCATTGAAAGGA 480
Qy 481 ATAGAACTGGCTGTGCGGGCCCACTCTGCCCCAGAAATACGAGGATGAACCATGACAGTT 540
Db 481 ATAGAACTGGCTGTGCGGGCCCACTCTGCCCCAGAAATACGAGGATGAACCATGACAGTT 540
Qy 541 TACCAGATCCCCATACACAGTGAACAGAGAGGGGAAGCACCAACCATGCGACAGTCCA 600
Db 541 TACCAGATCCCCATACACAGTGAACAGAGAGGGGAAGGCCAACCACTGGCAGAGTCCA 600
Qy 601 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATCTTTCAGACTCCGAGTCCGAATGAAAT 660
Db 601 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATCTTTCAGACTCCGAGTCCGAATGAAAT 660
Qy 661 GAGCCACACTTCCACATGGTGTAGCCAGAGAGGGGTGAGGACTCTTCCCTGGTC 720
Db 661 GAGCCACACTTCCACATGGTGTAGCCAGAGAGGGGTGAGGACTCTTCCCTGGTC 720
Qy 721 GTAGCTTTCATCTGTAGCTTCACTTAAAGAGAGAACTTCTTGGTGTCAAAGCAAG 780
Db 721 GTAGCTTTCATCTGTAGCTTCACTTAAAGAGAGAACTTCTTGGTGTCAAAGCAAG 780
Qy 781 GAGATGGGCTCCAGTGGGACAGCTGCCATCGCTCCCATCATTTGTGTGTCAAAGGAC 840
Db 781 GAGATGGGCTCCAGTGGGACAGCTGCCATCGCTCCCATCATTTGTGTGTCAAAGGAC 840
Qy 841 GGGAAAAAGCATCACTATGAAGGAAGAGATTTTGGCTGAAGAGCTGTGTACTCTTCCA 900
Db 841 GGGAAAAAGCATCACTATGAAGGAAGAGATTTTGGCTGAAGAGCTGTGTACTCTTCCA 900
Qy 901 GATCCTGGTCTGCTTTTGGTGGTGAATGTCCAGATGAAGCTTTCATTCAACCCATC 960
Db 901 GATCCTGGTCTGCTTTTGGTGGTGAATGTCCAGATGAAGCTTTCATTCAACCCATC 960
Qy 961 TGTGAGAAATGCCACTTTCAGAGGTACCAAGAAAGCAGATGCCCGCTGGCCCTTGGTG 1020
Db 961 TGTGAGAAATGCCACTTTCAGAGGTACCAAGAAAGCAGATGCCCGCTGGCCCTTGGTG 1020
Qy 1021 GTTCAATGCGCCCGACATCTGTCTTGTGGACAGCAGGTACACAGCAGTGTGATGAGAGG 1080
Db 1021 GTTCAATGCGCCCGACATCTGTCTTGTGGACAGCAGGTACACAGCAGTGTGATGAGAGG 1080
Qy 1081 TTTGGGCTGACACCCAGACCTTGGTCTGAATGAGAACTGTGCTCAGTTTCAACCTT 1140
Db 1081 TTTGGGCTGACACCCAGACCTTGGTCTGAATGAGAACTGTGCTCAGTTTCAACCTT 1140
Qy 1141 CGCAGCCACAAGATTCAAACCCAGCTCAACTCAACCCGGACATCTTCCCCCTGCTC 1200
Db 1141 CGCAGCCACAAGATTCAAACCCAGCTCAACTCAACCCGGACATCTTCCCCCTGCTC 1200
Qy 1201 ACCAGTTTCCGCTGTAAAGAGAGGGGCCCACTCTAGTGTGCCCTGTTTCAAGGTGAA 1260
Db 1201 ACCAGTTTCCGCTGTAAAGAGAGGGGCCCACTCTAGTGTGCCCTGTTTCAAGGTGAA 1260
Qy 1261 TGCCTCTCAAGTACCACTCCGTCCTGAGGAGGGGTGGCAGAGGGATGCCATTATTACT 1320
Db 1261 TGCCTCTCAAGTACCACTCCGTCCTGAGGAGGGGTGGCAGAGGGATGCCATTATTACT 1320
Qy 1321 TGCAATCTCAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCCAGCAGAGCGTG 1380
Db 1321 TGCAATCTCAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCCAGCAGAGCGTG 1380
Qy 1381 CAGGAGTACAGGAGGAGTGCAGGACCGGCCAGCCCGCAGCAGAGAAAGAGTCAAGTAC 1440
Db 1381 CAGGAGTACAGGAGGAGTGCAGGACCGGCCAGCCCGCAGCAGAGAAAGAGTCAAGTAC 1440
Qy 1441 CCAGAAATCATCTTCTTGGNAACAGGGTCTGCCATCCCGATCCGAGATTCGAAATGTCA 1500


```
Db 1441 CCAGAAATCATCTTCTTGGAAACAGGCTGCCATCCCATGAATTCGAATGTCAAT 1500
Qy 1501 GCCACACTTGTCAAATAGCCGCCACACAGCTCTCTGCTACTGCACTGTGGTGAAGGCACA 1560
Db 1501 GCCACACTTGTCAAATAGCCGCCACACAGCTCTCTGCTACTGCACTGTGGTGAAGGCACG 1560
Qy 1561 TTTGGGAGCTGTGGCTCATTAACGAGACACAGGTGGAGAGGTCCTGGGCAACCTGTGCT 1620
Db 1561 TTTGGGAGCTGTGGCTCATTAACGAGACACAGGTGGAGAGGTCCTGGGCAACCTGTGCT 1620
Qy 1621 GCTGTGTTGTCTCCACTGACCCAGATCACACAGGGCTCCCAAGTATCTTGCTG 1680
Db 1621 GCTGTGTTGTCTCCACTGACCCAGATCACACAGGGCTCCCAAGTATCTTGCTG 1680
Qy 1681 CAGAGAGAACGGCCCTTGGCATCTTTGGGAAAGCCGCTTACCCCTTTGCTGTGTGTGCC 1740
Db 1681 CAGAGAGAACGGCCCTTGGCATCTTTGGGAAAGCCCTTACCCCTTTGCTGTGTGTGCC 1740
Qy 1741 CCCAACAGCTCAAGCTGTGCTCAGCAGTACCAACACAGTGCAGAGTCTCTGCAC 1800
Db 1741 CCCAACAGCTCAAGCTGTGCTCAGCAGTACCAACACAGTGCAGAGTCTCTGCAC 1800
Qy 1801 CACATCAGTATGATTCCTGCCAATGCCCTTACGAGGGCTGAGATCTCCAGTCTGCA 1860
Db 1801 CACATCAGTATGATTCCTGCCAATGCCCTTACGAGGGCTGAGATCTCCAGTCTGCA 1860
Qy 1861 GTGGAAGATTCATCAGTTCGCTGTTCGCAATGTATTTGGAAGAGTTTCAGACCTGT 1920
Db 1861 GTGGAAGATTCATCAGTTCGCTGTTCGCAATGTATTTGGAAGAGTTTCAGACCTGT 1920
Qy 1921 CTGGTGGCAGCTCAAGCATCGTTTGGCTGTGGCTGTGCATCACACCTCTGGCTGAAA 1980
Db 1921 CTGGTGGCAGCTCAAGCATCGTTTGGCTGTGGCTGTGCATCACACCTCTGGCTGAAA 1980
Qy 1981 GTGGTCTATTCGGGGACACCTGCTCGAGGCTCTGGTCCGATGGGGAAGATGCC 2040
Db 1981 GTGGTCTATTCGGGGACACCTGCTCGAGGCTCTGGTCCGATGGGGAAGATGCC 2040
Qy 2041 ACCCTCTGATACATGAAGCCACCTCGAAGATGTTTGAAGAGAGCAAGTGGAAAG 2100
Db 2041 ACCCTCTGATACATGAAGCCACCTCGAAGATGTTTGAAGAGAGCAAGTGGAAAG 2100
Qy 2101 ACACACAGCACAAAGTCCAAAGCATACAGTGGGATCGGATGAACCGGAGTTCATT 2160
Db 2101 ACACACAGCACAAAGTCCAAAGCATACAGTGGGATCGGATGAACCGGAGTTCATT 2160
Qy 2161 ATGCTGAACCACTTCAGCCAGGCTATGCCAAGTCCCTCTTCAGCCCCAATTCAGC 2220
Db 2161 ATGCTGAACCACTTCAGCCAGGCTATGCCAAGTCCCTCTTCAGCCCCAATTCAC 2220
Qy 2221 GAGAAAGTGGGAGTTCCTTTGACCAATGAAGTCTGCTTTGGAGACTTTCACAAATG 2280
Db 2221 GAGAAAGTGGGAGTTCCTTTGACCAATGAAGTCTGCTTTGGAGACTTTCACAAATG 2280
Qy 2281 CCCAAGCTGATTCCTCCACTGAAGCCCTGTTTGTGGGACATCAGAGAGATGGAGGAG 2340
Db 2281 CCCAAGCTGATTCCTCCACTGAAGCCCTGTTTGTGGGACATCAGAGAGATGGAGGAG 2340
Qy 2341 CGCAGGAGAGCGGAGCTGGGAGAGTGGGAGGTCGGGGCGGCCCTCTGTCCAGGAGCTGGCA 2400
Db 2341 CGCAGGAGAGCGGAGCTGGGAGAGTGGGAGGTCGGGGCGGCCCTCTGTCCAGGAGCTGGCA 2400
Qy 2401 GCGGCTCGAGAGTGGGAGCTCAGCAGAGAGCGGGCCACACAGAGAGGCCACAGGCC 2460
Db 2401 GCGGCTCGAGAGTGGGAGCTCAGCAGAGAGCGGGCCACACAGAGAGGCCACAGGCC 2460
Qy 2461 AAGAAGTTCAGAGCCCAAGTGA 2481
Db 2461 AAGAAGTTCAGAGCCCAAGTGA 2481
```

RESULT 8

```
AK094687
LOCUS Homo sapiens cDNA FLJ37368 fis, clone BRAMY2024530, highly similar
DEFINITION to Homo sapiens prostate cancer susceptibility protein HPC2/ELAC2
mRNA.
ACCESSION AK094687
VERSION AK094687.1 GI:21753794
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens amygdala cDNA to mRNA, clone_lib:BRAMY2
clone:BRAMY2024530.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Tashiro,H., Yanazaki,M., Watanabe,K., Kumagai,A., Itakura,S.,
Fukuzumi,Y., Fujimori,Y., Komiyama,M., Sugiyama,T., Irie,R.,
Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,
Oshina,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,
Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2907)
Isogai,T. and Yamamoto,J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
Location/Qualifiers
source 1..2907
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="BRAMY2024530"
/tissue_type="amygdala"
/clone_lib="BRAMY2"
/note="cloning vector: pME18SFL3"
BASE COUNT 696 a 788 c 839 g 584 t
ORIGIN
```

```
Query Match 94.7%; Score 2349; DB 9; Length 2907;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 2419; Conservative 0; Mismatches 5; Indels 57; Gaps 1;
Qy 1 ATGTGGGGCGCTTTGCTCGCTGCTGCGGTCCGGCCGCGACGCCACCATCTCGCAGGAGCG 60
Db 59 ATGTGGGGCGCTTTGCTCGCTGCTGCGGTCCGGCCGCGACGCCACCATCTCGCAGGAGCG 118
Qy 61 ACCATATCGCAGGACCGCCCGCGCGAGCGCGCGCAAGGACCCCGCTCGGGCACCTG 120
Db 119 ACCATATCGCAGGACCGCCCGCGCGAGCGCGCGCAAGGACCCCGCTCGGGCACCTG 178
Qy 121 CGCAGCGAGAGAGCGCGGACCGTCTCGGGGTGCTCCGGCGGCCCAACACCGTGTACCTG 180
Db 179 CGCAGCGAGAGAGCGCGGACCGTCTCGGGGTGCTCCGGCGGCCCAACACCGTGTACCTG 238
Qy 181 CAGTGTGTGGCAGCGGCTAGCGGGACTCTGGGCGCGCGCTCTACGTTCTTCTCCAGTTC 240
Db 239 CAGTGTGTGGCAGCGGCTAGCGGGACTCTG----- 268
Qy 241 AACCGGTATCTTTCAACTGTGGAGAGCGTTTCAGAGACTCATCGAGGACACAAGTTA 300
Db 269 -----GCGGTTTCAGAGACTCATCGAGGACACAAGTTA 301
```

QY 301 AAGTTGCTCGCTGGCAACAATATTTCTTGACAGAAATGCACTGGTCTAAATGTTGGGGGC 360
Db |||||
QY 302 AAGTTGCTCGCTGGCAACAATATTTCTTGACAGAAATGCACTGGTCTAAATGTTGGGGGC 361
Db |||||
QY 361 TTAAGTGAATGATTTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGCTACTTTCTCTGGA 420
Db |||||
QY 362 TTAAGTGAATGATTTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGCTACTTTCTCTGGA 421
Db |||||
QY 421 CCTCCAACTCGGAAAAATACCTCGAAGCAATCAAAAATATTTTCTGGTCCATTTGAAAGGA 480
Db |||||
QY 422 CCTCCAACTCGGAAAAATACCTCGAAGCAATCAAAAATATTTTCTGGTCCATTTGAAAGGA 481
Db |||||
QY 481 ATAGAACTGGCTGCGGGCCCACTCTGCCCAGAAATACGAGGATGAACCAATGACAGTT 540
Db |||||
QY 482 ATAGAACTGGCTGCGGGCCCACTCTGCCCAGAAATACGAGGATGAACCAATGACAGTT 541
Db |||||
QY 541 TACCAGATCCCCATACACAGTGAACAGAGGGGAAAGCACAACCACTGGCAGAGTCCA 600
Db |||||
QY 542 TACCAGATCCCCATACACAGTGAACAGAGGGGAAAGCACAACCACTGGCAGAGTCCA 601
Db |||||
QY 601 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCATCTTCAGACTCCGAGTCCGAATGAAAT 660
Db |||||
QY 602 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCATCTTCAGACTCCGAGTCCGAATGAAAT 661
Db |||||
QY 661 GAGCCACACCTTCCACATGTTTAAAGAGAGAAACCTCTTGGTGTCAAAGCAAG 780
Db |||||
QY 721 GTAGCTTTTCACTGTAAGCTTCACTTAAAGAGAGAAACCTCTTGGTGTCAAAGCAAG 780
Db |||||
QY 722 GTAGCTTTTCACTGTAAGCTTCACTTAAAGAGAGAAACCTCTTGGTGTCAAAGCAAG 781
Db |||||
QY 781 GAGATGGGCTCCAGTGGGACAGTCCATCGCTCCCATCAATGCTGTCAAGAAC 840
Db |||||
QY 782 GAGATGGGCTCCAGTGGGACAGTCCATCGCTCCCATCAATGCTGTCAAGAAC 841
Db |||||
QY 841 GGGAAAGCATCACTCATGAAGGAGAGAGATTTTGGCTGAAGAGCTGTACTCCTCCA 900
Db |||||
QY 842 GGGAAAGCATCACTCATGAAGGAGAGAGATTTTGGCTGAAGAGCTGTACTCCTCCA 901
Db |||||
QY 901 GATCTGTGTGCTTTTGTGTGTGTAGAAATGTCAGAGTGAAGAGCTTCAATCAACCCATC 960
Db |||||
QY 902 GATCTGTGTGCTTTTGTGTGTGTAGAAATGTCAGAGTGAAGAGCTTCAATCAACCCATC 961
Db |||||
QY 961 TGTGAATGCACTTTTCAAGGTACCAAGAAAGGAGATGCCCGCTGGCTTGGTG 1020
Db |||||
QY 962 TGTGAATGCACTTTTCAAGGTACCAAGAAAGGAGATGCCCGCTGGCTTGGTG 1021
Db |||||
QY 1021 GTTCATGCGCCCGCAGCATCTGTGCTTGTGGACAGCAGGTACCCAGCTGGATGGAGAGG 1080
Db |||||
QY 1081 TTTTGGGCTTGACACCCAGCACTTGTGCTCTGAATGAGAACTGTGCTCTCAATGTCAGTTCACAACTT 1140
Db |||||
QY 1082 TTTTGGGCTTGACACCCAGCACTTGTGCTCTGAATGAGAACTGTGCTCTCAATGTCAGTTCACAACTT 1141
Db |||||
QY 1141 CGCAGCCAAAGATTCAAACCCAGTCAACCTCACTCCAGCCGGAATCTTCCCTGCTGTC 1200
Db |||||
QY 1142 CGCAGCCAAAGATTCAAACCCAGTCAACCTCACTCCAGCCGGAATCTTCCCTGCTGTC 1201
Db |||||
QY 1201 ACCAGTTTCCGCTGTAAAGAGAGGGCCCACTCTGAGTGCCCATGGTTCAGGGTCAA 1260
Db |||||
QY 1261 TGCCTCTCAAGTACAGCTCCGCTCCAGAGGGAGTGGCAGAGGGATGCCATTATTACT 1320
Db |||||
QY 1262 TGCCTCTCAAGTACAGCTCCGCTCCAGAGGGAGTGGCAGAGGGATGCCATTATTACT 1321
Db |||||
QY 1321 TGCATCTGTAGGAATTCATAGTTGAGGGCTGTGAGCTTCCCACTTCAGCAGAGCGTG 1380
Db |||||
QY 1322 TGCATCTGTAGGAATTCATAGTTGAGGGCTGTGAGCTTCCCACTTCAGCAGAGCGTG 1381
Db |||||
QY 1381 CAGGAGTACAGGAGGTGCGCAGGACGGCCCGCCAGCAGAGAGAAAGTACAGTAC 1440
Db |||||

Db 1382 CAGAGTACAGGAGGTGCGCAGACGGCCAGCCCGCAGAGAGAAAGTACAGTAC 1441
QY |||||
QY 1441 CCAGAAATCATTTCTTTGGAAACAGGCTGTCCTATCCCGATGAAGATTCGAAATGTCAGT 1500
Db |||||
QY 1442 CCAGAAATCATTTCTTTGGAAACAGGCTGTCCTATCCCGATGAAGATTCGAAATGTCAGT 1501
Db |||||
QY 1501 GCCACACTTGTCAACATTAAGCCCGACACAGTCTCTGCTACTGGAATGTGGTGGAGGCACA 1560
Db |||||
QY 1502 GCCACACTTGTCAACATTAAGCCCGACACAGTCTCTGCTACTGGAATGTGGTGGAGGCACG 1561
Db |||||
QY 1561 TTTGGGAGAGTGTCCCTCATTAACGAGAGACAGGTGACAGGGTCTTGGGACACCTTGCT 1620
Db |||||
QY 1562 TTTGGGAGAGTGTCCCTCATTAACGAGAGACAGGTGACAGGGTCTTGGGACACCTTGCT 1621
Db |||||
QY 1621 GCTGTGTTTGTGTCCTCCACCTGACAGCATCAACACAGGGCTTGCACAGTATCTTGCTG 1680
Db |||||
QY 1622 GCTGTGTTTGTGTCCTCCACCTGACAGCATCAACACAGGGCTTGCACAGTATCTTGCTG 1681
Db |||||
QY 1681 CAGAGAAACCGGCTTTGGCATCTTTGGGAAAGCGCTTCAACCTTTGCTGGTGGTGGCC 1740
Db |||||
QY 1682 CAGAGAAACCGGCTTTGGCATCTTTGGGAAAGCGCTTCAACCTTTGCTGGTGGTGGCC 1741
Db |||||
QY 1741 CCCAACAGCTCAAAGCTGSGCTCCAGCAGTACCAGAACAGTCCAGGAGGTCTCTGCAC 1800
Db |||||
QY 1742 CCCAACAGCTCAAAGCTGSGCTCCAGCAGTACCAGAACAGTCCAGGAGGTCTCTGCAC 1801
Db |||||
QY 1801 CACATCAGTATGATTTCTGTCCTCAAATGCTTTAGGAGGGCTGAGATCTCCAGTCTTGCA 1860
Db |||||
QY 1802 CACATCAGTATGATTTCTGTCCTCAAATGCTTTAGGAGGGCTGAGATCTCCAGTCTTGCA 1861
Db |||||
QY 1861 GTGAAAGATGATCAGTTGCTGTTGCGAACATGATGATTTGGAAGAGTTCAGACCTGT 1920
Db |||||
QY 1862 GTGAAAGATGATCAGTTGCTGTTGCGAACATGATGATTTGGAAGAGTTCAGACCTGT 1921
Db |||||
QY 1921 CTGCTGCGGACCTCAAGCATGCGTTTGGCTGTCGCTGTTGTCACACTCTGCTGGTGGAAA 1980
Db |||||
QY 1922 CTGCTGCGGACCTCAAGCATGCGTTTGGCTGTCGCTGTTGTCACACTCTGCTGGTGGAAA 1981
Db |||||
QY 1981 GTGCTCTATTTCCGGGACACCATGCTCCCTGCGAGCTCTGCTGCGGATGGGAAAGATGCC 2040
Db |||||
QY 1982 GTGCTCTATTTCCGGGACACCATGCTCCCTGCGAGCTCTGCTGCGGATGGGAAAGATGCC 2041
Db |||||
QY 2041 ACCCTCTGATACATGAAGCCACCTCGAAGATGGTTTGGAGAGGAGCAGTGGAAAG 2100
Db |||||
QY 2042 ACCCTCTGATACATGAAGCCACCTCGAAGATGGTTTGGAGAGGAGCAGTGGAAAG 2101
Db |||||
QY 2101 ACACAGCACAACGCTCCCAAGCCTCAGCGTGGGATGCGGATGAACGCGAGTTCATT 2160
Db |||||
QY 2102 ACACAGCACAACGCTCCCAAGCCTCAGCGTGGGATGCGGATGAACGCGAGTTCATT 2161
Db |||||
QY 2161 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGTCCCTCTTTTCAGCCCCCACTTCAGC 2220
Db |||||
QY 2162 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGTCCCTCTTTTCAGCCCCCACTTCAGC 2221
Db |||||
QY 2221 GAGAAAGTGGAGTGGCTTTGACCAATGAAGTCTGCTTTGAGAGCTTTCCAAACATG 2280
Db |||||
QY 2222 GAGAAAGTGGAGTGGCTTTGACCAATGAAGTCTGCTTTGAGAGCTTTCCAAACATG 2281
Db |||||
QY 2281 CCCAAGCTGATTTCCCACTGAAAGCCTGTTTGTGCGGACATCGAGGAGATGGAGGAG 2340
Db |||||
QY 2282 CCCAAGCTGATTTCCCACTGAAAGCCTGTTTGTGCGGACATCGAGGAGATGGAGGAG 2341
Db |||||
QY 2341 CGCAGGAGAAAGCGGAGCTGCGGAGGTGCGGCGGCTCTCTGTCCAGGAGCTGGCA 2400
Db |||||
QY 2342 CGCAGGAGAAAGCGGAGCTGCGGAGGTGCGGCGGCTCTCTGTCCAGGAGCTGGCA 2401
Db |||||
QY 2401 GCGGCTTGGAGATGGGAGCTCAGCAGAGCGGCGGACACACAGAGGAGCCACAGGCC 2460
Db |||||
QY 2402 GCGGCTTGGAGATGGGAGCTCAGCAGAGCGGCGGACACACAGAGGAGCCACAGGCC 2461
Db |||||
QY 2461 AAGAAAGTTCAGAGCCAGTGA 2481
Db |||||

Db 2462 AAGAAGTCTAGAGCCCACTGA 2482

RESULT 9
AK094333
LOCUS
DEFINITION
Homo sapiens cDNA FLJ37014 fis, clone BRACE2010203, highly similar to Homo sapiens prostate cancer susceptibility protein HPC2/ELAC2 mRNA.
AK094333 2734 bp mRNA linear PRI 15-JUL-2002
Homo sapiens cDNA FLJ37014 fis, clone BRACE2010203, highly similar to Homo sapiens prostate cancer susceptibility protein HPC2/ELAC2 mRNA.
AK094333 1 GI:21753374
oligo capping; fis (full insert sequence).
Homo sapiens cerebellum cDNA to mRNA, clone_lib:BRACE2
clone:BRACE2010203.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Nishi,T., Ota,T., Nakagawa,S., Senoh,A., Mizuguchi,H., Inagaki,H., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahara,K., Masuno,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2734)
Isogai,T. and Yamamoto,J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB.
RAB; annotation: HRI and RAB.
Location/Qualifiers
1. .2734
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="BRACE2010203"
/tissue_type="cerebellum"
/clone_lib="BRACE2"
/note="cloning vector: pME18SFL3"
BASE COUNT 671 a 708 c 724 g 631 t
ORIGIN
Query Match 70.3%; Score 1745; DB 9; Length 2734;
Best Local Similarity 100.0%; Pred. No.: 0;
Matches 1745; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 737 AGCTTCACCTTAAGAGAGGAAACTCTTTGGTGCTCAAGCAAGAGATGGCCCTCCAG 796
Db 565 AGCTTCACCTTAAGAGAGGAAACTCTTTGGTGCTCAAGCAAGAGATGGCCCTCCAG 624
Qy 797 TTGGACAGCTGCCATCGCTCCCATCATTTGTGTGTCAAGGACGGGAAAGCATCACTC 856
Db 625 TTGGACAGCTGCCATCGCTCCCATCATTTGTGTGTCAAGGACGGGAAAGCATCACTC 684
Qy 857 ATGAGGAGAGAGATTTTGGCTGAGAGCTGTGTACTCTCCAGATCCTGTGCTGCTT 916
Db 685 ATGAAGGAAGAGAGATTTTGGCTGAGAGCTGTGTACTCTCCAGATCCTGTGCTGCTT 744
Qy 917 TTGTGGTGGTAGAATGTCCAGATGAAGCTTCATTCAACCCCATCTGTGAGAAATGCCACT 976
Db 745 TTGTGGTGGTAGAATGTCCAGATGAAGCTTCATTCAACCCCATCTGTGAGAAATGCCACT 804
Qy 977 TTCAGAGTACCAAGGAAAGGAGATGCCCGCTGGCTTGGTGTTCACATGCCCCCAG 1036

Db 1885 AAGCCACCTCGAAGATGGTTTGAAGAGGAGCAGTGGAAAAGACACACAGCAACAGT 1944

Qy 2117 CCCAAGCCATCAGCTGGGATGCGGATGAACGCGGAGTTCAATATGCTGAACCACTTCA 2176

Db 1945 CCCAAGCCATCAGCTGGGATGCGGATGAACGCGGAGTTCAATATGCTGAACCACTTCA 2004

Qy 2177 GCCAGCGTATGCCAAGGTCCTCTTCAGCCCCCACTTCAGCGAGAAAGTGGGAGTTG 2236

Db 2005 GCCAGCGTATGCCAAGGTCCTCTTCAGCCCCCACTTCAGCGAGAAAGTGGGAGTTG 2064

Qy 2237 CTTTGACCAATGAAGTCTGCTTTGGAGACTTTCACCAATGCCCAGCTGATTCCTCC 2296

Db 2065 CTTTGACCAATGAAGTCTGCTTTGGAGACTTTCACCAATGCCCAGCTGATTCCTCC 2124

Qy 2297 CACTGAAGCCCTGTTTCTGCGCATCGAGAGATGGAGAGCGCAGGAGAAAGCGG 2356

Db 2125 CACTGAAGCCCTGTTTCTGCGCATCGAGAGATGGAGAGCGCAGGAGAAAGCGG 2184

Qy 2357 AGCTGGCAGTGGCGGCGCTCTCTGTCAGGAGCTGGCAGGCGCTGGAGGATG 2416

Db 2185 AGCTGGCAGTGGCGGCGCTCTCTGTCAGGAGCTGGCAGGCGCTGGAGGATG 2244

Qy 2417 GGGAGCCTCAGCAGAAAGCGGCGCCACACAGAGAGCCACAGGCCAAGAGTTCAGAGCCC 2476

Db 2245 GGGAGCCTCAGCAGAAAGCGGCGCCACACAGAGAGCCACAGGCCAAGAGTTCAGAGCCC 2304

Qy 2477 AGTGA 2481

Db 2305 AGTGA 2309

RESULT 10

AK074244

LOCUS

DEFINITION Homo sapiens cDNA FLJ23664 fis, clone HEP03495, highly similar to Putative prostate cancer susceptibility protein.

ACCESSION AK074244

VERSION AK074244.1 GI:18676795

KEYWORDS oligo capsing; fis (full insert sequence).

SOURCE Homo sapiens hepatoma cell_line:HepG2 cDNA to mRNA, clone_lib:HEP clone:HEP03495.

ORGANISM Homo sapiens

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Kawabata,A., Hikiiji,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.

TITLE NEDO human cDNA sequencing project

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 2163)

AUTHORS Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T., Shibahara,T., Tanaka,T. and Nakamura,Y.

TITLE Direct Submission

JOURNAL Submitted (14-FEB-2002) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)

COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing; Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).

FEATURES

source

1..2163

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="HEP03495"

/cell_line="HepG2"

/cell_type="hepatoma"

/clone_lib="HEP"

BASE COUNT 530 a 582 c 613 g 438 t

ORIGIN

Query Match 69.1%; Score 1715.2; DB 9; Length 2163;

Best Local Similarity 99.8%; Pred. No. 0;

Mismatches 1717; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 762 CTTGGTCTCTCAAGCAAGAGATGGGCTCCAGTTGGGACAGCTGCCATCGCTCCAT 821

Db 1 CTTGGTCTCTCAAGCAAGAGATGGGCTCCAGTTGGGACAGCTGCCATCGCTCCAT 60

Qy 822 CATTGCTGCTCAAGACCGGAAAGCATCACTCATGAAGGAGAGATTTTGGCTGA 881

Db 61 CATTGCTGCTCAAGACCGGAAAGCATCACTCATGAAGGAGAGATTTTGGCTGA 120

Qy 882 AGAGCTGTGTACTCTCTCCAGATCCTGGTGTGCTTTTGTGTGTAGATGTCCAGATGA 941

Db 121 AGAGCTGTGTACTCTCTCCAGATCCTGGTGTGCTTTTGTGTGTAGATGTCCAGATGA 180

Qy 942 AAGCTTCATTCAACCCATCTGTGAGAAATGCCACTTTTCAGAGGTACCAAGAAAGGAGA 1001

Db 181 AAGCTTCATTCAACCCATCTGTGAGAAATGCCACTTTTCAGAGGTACCAAGAAAGGAGA 240

Qy 1002 TGCCCCCGTGGCTTGGTGTTCACATGGCCCCAGACATCTGTCTTGTGGACACAGGTA 1061

Db 241 TGCCCCCGTGGCTTGGTGTTCACATGGCCCCAGACATCTGTCTTGTGGACACAGGTA 300

Qy 1062 CCAGCAGTGGATGGAGAGTTTGGGCTTGACACCCAGCAGCTTGGTCTCTGAATGAGAACTG 1121

Db 301 CCAGCAGTGGATGGAGAGTTTGGGCTTGACACCCAGCAGCTTGGTCTCTGAATGAGAACTG 360

Qy 1122 TGCCTCAGTTCAACACCTTCGACGCCACAGATTCACAAACAGCAGCTCAACCTCATCCACC 1181

Db 361 TGCCTCAGTTCAACACCTTCGACGCCACAGATTCACAAACAGCAGCTCAACCTCATCCACC 420

Qy 1182 GGACATCTTCCCTCTGCTCACCAGTTTCCGCTGTGAAGAGAGGCGCCACCTCAGTGT 1241

Db 421 GGACATCTTCCCTCTGCTCACCAGTTTCCGCTGTGAAGAGAGGCGCCACCTCAGTGT 480

Qy 1242 GCCCATGTTTCAGGGTGAATCCCTCTCAAGTACACAGCTCCGTCCTCCAGGAGGAGTGGA 1301

Db 481 GCCCATGTTTCAGGGTGAATCCCTCTCAAGTACACAGCTCCGTCCTCCAGGAGGAGTGGA 540

Qy 1302 GAGGATGCCATTAATTAATCTTGCATCTCTGAGGAATTCATAGTTGAGCGCTGAGCTTCC 1361

Db 541 GAGGATGCCATTAATTAATCTTGCATCTCTGAGGAATTCATAGTTGAGCGCTGAGCTTCC 600

Qy 1362 CAACCTTCCAGCAGCGCTGCAGGAGTACAGGAGAGTCCGACGAGCGCCAGCCCGCCAGC 1421

Db 601 CAACCTTCCAGCAGCGCTGCAGGAGTACAGGAGAGTCCGACGAGCGCCAGCCCGCCAGC 660

Qy 1422 AGAGAAAGAGTCACTACCCAGAAATCACTTCTTCTTGGAAACAGGGTCTGCTATCCCGAT 1481

Db 661 AGAGAAAGAGTCACTACCCAGAAATCACTTCTTCTTGGAAACAGGGTCTGCTATCCCGAT 720

Qy 1482 GAAGATTGCAAAATGCTAGTGCCACATTTGCAACATGAAGCCCCACAGCTCTGCTACT 1541

Db 721 GAAGATTGCAAAATGCTAGTGCCACATTTGCAACATGAAGCCCCACAGCTCTGCTACT 780

Qy 1542 GGACTGTGGTGAAGGACATTTGGGAGCTGTGCGCTCATTTACGGAGACAGGTGGACAG 1601

Db 781 GGACTGTGGTGAAGGACATTTGGGAGCTGTGCGCTCATTTACGGAGACAGGTGGACAG 840

Qy 1602 GGTCTCTGGCACCTTGGCTGT 1661

Db 841 GGTCTCTGGCACCTTGGCTGT 900

Qy 1662 CTTGCCAAGTATCTTGTGTGAGAGAGAACCGGCTCTGGCATCTTTGGAAAGCGCTTCA 1721

Db 901 CTTGCCAAGTATCTTGTGTGAGAGAGAACCGGCTCTGGCATCTTTGGAAAGCGCTTCA 960

Qy 1722 CCCTTTGT 1781

Query Match	33.2%; Score 823; DB 9; Length 2793;
Best Local Similarity	100.0%; Pred. No. 7.1e-181;
Matches	823; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1659	GGGCTTGCCCAAGATCTTGTGCAGAGAGAACGCGCTTGGCATCTTTGGGAAGCCGCT	1718
Db	1546	GGGCTTGCCCAAGATCTTGTGCAGAGAGAACGCGCTTGGCATCTTTGGGAAGCCGCT	1605
Qy	1719	TCACCCCTTTGCTGGTGGTTGCCCAACACAGCTCAAAAGCCCTGGCTCCAGCAGTACCACAA	1778
Db	1606	TCACCCCTTTGCTGGTGGTTGCCCAACACAGCTCAAAAGCCCTGGCTCCAGCAGTACCACAA	1665
Qy	1779	CCAGTGCAGAGAGTCTGCACACATCAGTATGATTCTTGCCAAATGCCCTTCAGGAAGG	1838
Db	1666	CCAGTGCAGAGAGTCTGCACACATCAGTATGATTCTTGCCAAATGCCCTTCAGGAAGG	1725
Qy	1839	GGCTGAGATCTCCAGTCCCTGCAGTGGAAAGATTGATCAGTTTCGCTTGGCAATGTGA	1898
Db	1726	GGCTGAGATCTCCAGTCCCTGCAGTGGAAAGATTGATCAGTTTCGCTTGGCAATGTGA	1785
Qy	1899	TTTGGAAAGAGTTTCAGACCTCTGTGGTGGCGCACTGCAAGCATGCGTTTGGCTCTCGCT	1958
Db	1786	TTTGGAAAGAGTTTCAGACCTCTGTGGTGGCGCACTGCAAGCATGCGTTTGGCTCTCGCT	1845
Qy	1959	GGTGCAACCTCTGGCTGGAAAGTGTCTATTCCGGGGACACCATGCCCTGCGAGGCTCT	2018
Db	1846	GGTGCAACCTCTGGCTGGAAAGTGTCTATTCCGGGGACACCATGCCCTGCGAGGCTCT	1905
Qy	2019	GGTCCGAGTGGGNAAGATGCCACCTCTCTGATACATGAAGCCACCTCGGAAGTGGTTT	2078
Db	1906	GGTCCGAGTGGGNAAGATGCCACCTCTCTGATACATGAAGCCACCTCGGAAGTGGTTT	1965
Qy	2079	GGAAAGAGAAAGCAGTGGAAAGAGACACACAGCAACCTGCCAAGCCATCAGCGTGGGGAT	2138
Db	1966	GGAAAGAGAAAGCAGTGGAAAGAGACACACAGCAACCTGCCAAGCCATCAGCGTGGGGAT	2025
Qy	2139	GCAGATGAACCGAGTTCATTATGTGTGAACCATCTTCAGCCAGCGCTATGCCAAGGTCCC	2198
Db	2026	GCAGATGAACCGAGTTCATTATGTGTGAACCATCTTCAGCCAGCGCTATGCCAAGGTCCC	2085
Qy	2199	CTCTTTCAGGCCCACTTCAGCAGAGAAAGTGGAGTTGCCCTTTGACACATGAAGTCTG	2258
Db	2086	CTCTTTCAGGCCCACTTCAGCAGAGAAAGTGGAGTTGCCCTTTGACACATGAAGTCTG	2145
Qy	2259	CTTTGGAGACTTTTCCAAACATGCCAAGCTGATTCCCCCACTGAAAGCCCTGTTTGTCTGG	2318
Db	2146	CTTTGGAGACTTTTCCAAACATGCCAAGCTGATTCCCCCACTGAAAGCCCTGTTTGTCTGG	2205
Qy	2319	CGACATCGAGGAGATGGAGAGCGCAGGGAGAACGCGGAGCTGCGCAGGTGCGGGCGGC	2378
Db	2206	CGACATCGAGGAGATGGAGAGCGCAGGGAGAACGCGGAGCTGCGCAGGTGCGGGCGGC	2265
Qy	2379	CTCTCTGTCAGGAGAGTGCAGCGCGCTTGGAGAGATGGGGAGCCTCAGCAGAAAGCGGGC	2438
Db	2266	CTCTCTGTCAGGAGAGTGCAGCGCGCTTGGAGAGATGGGGAGCCTCAGCAGAAAGCGGGC	2325
Qy	2439	CCACAGAGAGGCCACAGGCCCAAGAGGTCAGAGCCCACTGA	2481
Db	2326	CCACAGAGAGGCCACAGGCCCAAGAGGTCAGAGCCCACTGA	2368

RESULT 13

AC005277

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AC005277

Homo sapiens chromosome 17, clone hRPK.597_M.12, complete sequence.

AC005277

AC005277.1

GI:3337311

HTG.

Homo sapiens.

Homo sapiens

REFERENCE

1

(bases 1 to 118788)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

AUTHORS      Birren,B., Fasman,K., Linton,L., Nusbaum,C. and Lander,E.
TITLE        Homo sapiens chromosome 17, clone hRPK.597_M.12
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 118788)
AUTHORS      Birren,B., Fasman,K., Linton,L., Nusbaum,C., Lander,E., Allen,N.,
            Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Boatn,C.,
            Boutwell,C., Brown,A., Castle,A., Cerny,J., Cooke,P., Depayre,E.,
            Devon,K., Dewar,K., Donelan,L., Etemadi,S., Ferreira,P.,
            FitzHugh,W., Forrest,C., Funke,R., Gage,D., Gardyna,S.,
            Gensheimer,S., Geraghty,K., Gilmartin,T., Grant,G., Hagos,B.,
            Harris,K., Horton,L., Howland,J.C., Hui,L., Jacotot,L., Kann,L.,
            Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
            Meldrim,J., Mollia,M., Morris,W., Morrow,J., Mychaleckyj,J.,
            Nachman,A., Nahf,R., Naylor,J., Niloff,M., O'Connor,T., Pavlin,B.,
            Peterson,K., Riley,R., Roberts,D., Rossello,R., Roy,A., Shyam,R.,
            Stange-Thomann,N., Stilwell,J., Stojanovic,N., Stone,C.,
            Strickland,C., Subramanian,A., Torruella-Miller,I., Vassiliev,H.,
            Vo,A., Wagner,A., Wang,B., Wheeler,J., Wu,Y., Ye,W.J., Zhao,J. and
            Zody,M.
TITLE        Direct Submission
JOURNAL      Submitted (10-JUL-1998) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE    3 (bases 1 to 118788)
AUTHORS      Birren,B., Fasman,K., Linton,L., Nusbaum,C., Lander,E., Allen,N.,
            Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Boatn,C.,
            Boutwell,C., Brown,A., Castle,A., Cerny,J., Cooke,P., Depayre,E.,
            Devon,K., Dewar,K., Donelan,L., Etemadi,S., Ferreira,P.,
            FitzHugh,W., Forrest,C., Funke,R., Gage,D., Gardyna,S.,
            Gensheimer,S., Geraghty,K., Gilmartin,T., Grant,G., Hagos,B.,
            Harris,K., Horton,L., Howland,J.C., Hui,L., Jacotot,L., Kann,L.,
            Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
            Meldrim,J., Mollia,M., Morris,W., Morrow,J., Mychaleckyj,J.,
            Nachman,A., Nahf,R., Naylor,J., Niloff,M., O'Connor,T., Pavlin,B.,
            Peterson,K., Riley,R., Roberts,D., Rossello,R., Roy,A., Shyam,R.,
            Stange-Thomann,N., Stilwell,J., Stojanovic,N., Stone,C.,
            Strickland,C., Subramanian,A., Torruella-Miller,I., Vassiliev,H.,
            Vo,A., Wagner,A., Wang,B., Wheeler,J., Wu,Y., Ye,W.J., Zhao,J. and
            Zody,M.
TITLE        Direct Submission
JOURNAL      Submitted (23-JUL-1998) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT      On Jul 23, 1998 this sequence version replaced GI:3335015.
            All repeats were identified using RepeatMasker: Smit, A.F.A. &
            Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html

FEATURES     source
            Location/Qualifiers
            1..118788
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="17"
            /map="17"
            /clone="hRPK.597_M.12"
            /clone_lib="RC1-11 human BAC library"
            795..884
            /rpt_family="MIR"
            complement(932..1000)
            /rpt_family="MIR"
            2044..2076
            /rpt_family="AT_rich"
            2237..2713
            /rpt_family="MER34"
            2718..3477
            /rpt_family="L2"
            complement(3478..3821)
            /rpt_family="MLT1A"
            3822..4007
            /rpt_family="L2"
            complement(4015..4281)
            /rpt_family="MER89"

repeat_region complement(4282..4321)
repeat_region /rpt_family="(TAA)n"
repeat_region complement(4328..4608)
repeat_region /rpt_family="AluSx"
repeat_region complement(4609..4895)
repeat_region /rpt_family="MER89"
repeat_region complement(4896..5186)
repeat_region /rpt_family="AluSp"
repeat_region 5219..5479
repeat_region /rpt_family="L2"
repeat_region complement(5481..5740)
repeat_region /rpt_family="MER34"
repeat_region complement(5963..6275)
repeat_region /rpt_family="AluSp"
repeat_region complement(6411..6494)
repeat_region /rpt_family="MER5A"
repeat_region 6830..7013
repeat_region /rpt_family="MLT1C"
repeat_region 7014..7318
repeat_region /rpt_family="AluYb8"
repeat_region 7319..7581
repeat_region /rpt_family="MLT1C"
repeat_region 7721..8032
repeat_region /rpt_family="AluJb"
repeat_region 8039..8333
repeat_region /rpt_family="AluSx"
repeat_region 9431..9460
repeat_region /rpt_family="AT_rich"
repeat_region 9662..9829
repeat_region /rpt_family="MIR"
repeat_region complement(11302..11728)
repeat_region /rpt_family="MER39"
repeat_region complement(11824..11884)
repeat_region /rpt_family="LTR16C"
repeat_region complement(11928..12108)
repeat_region /rpt_family="LTR16C"
repeat_region 12453..12725
repeat_region /rpt_family="AluJb"
repeat_region 13377..13785
repeat_region /rpt_family="MSTB"
repeat_region 14118..14153
repeat_region /rpt_family="(CA)n"
repeat_region complement(15320..15620)
repeat_region /rpt_family="AluJo"
repeat_region 15883..15759
repeat_region /rpt_family="(CA)n"
repeat_region complement(15838..15893)
repeat_region /rpt_family="MER91A"
repeat_region complement(15915..16080)
repeat_region /rpt_family="MER91A"
repeat_region complement(16474..16544)
repeat_region /rpt_family="L2"
repeat_region 17907..17920
repeat_region /rpt_family="L1ME"
repeat_region 18271..18311
repeat_region /rpt_family="L2"
repeat_region complement(18368..18529)
repeat_region /rpt_family="AluJb"
repeat_region complement(18530..18560)
repeat_region /rpt_family="(CA)n"
repeat_region complement(18561..18681)
repeat_region /rpt_family="AluJb"
repeat_region 18760..19065
repeat_region /rpt_family="AluSg"
repeat_region complement(19638..19896)
repeat_region /rpt_family="MIR"
repeat_region complement(19971..20260)
repeat_region /rpt_family="L1ME3A"
repeat_region 20263..20552
repeat_region /rpt_family="MER58B"
repeat_region complement(20553..20816)
repeat_region /rpt_family="L1ME3A"
repeat_region complement(20821..21069)

```

```

/rpt_family="LiM4"
complement(21071..21100)
/rpt_family="TAAA)n"
complement(21101..21138)
/rpt_family="AluSx"
complement(21348..21598)
/rpt_family="LiM4"
21799..22104
/rpt_family="AluGo"
22151..22562
/rpt_family="MSTA"
23978..24287
/rpt_family="AluSq"
24852..25199
/rpt_family="L2"
complement(25697..25985)
/rpt_family="MLT1D"
25816..25895
/notes="Single-stranded coverage."
complement(25994..26099)
/rpt_family="purine-rich"
complement(26122..26291)
/rpt_family="MLT1D"
26312..26337
/rpt_family="CAAAA)n"
26456..26480
/rpt_family="GGGAA)n"
complement(26576..26781)
/rpt_family="WER6B"
complement(26912..26990)
/rpt_family="WER21B"
26991..27290
/rpt_family="AluSx"
complement(27291..27975)
/rpt_family="WER21B"
28034..28197
/rpt_family="LTR28"
28266..28795
/rpt_family="LTR28"
28843..28916
/rpt_family="LTR28"
28920..29005
/rpt_family="MIR"
29140..29233
repeat_region

```

Query Match	10.0%	Score 247.4;	DB 9;	Length 118788;
Best Local Similarity	97.7%;	Pred. No. 1.1e-46;		
Matches 251;	Conservative 0;	Mismatches 6;	Indels 0;	Gaps 0;
QY	1	ATGTGGGGCTTTTCTCTGCTCTGGGTCCGGGCGGACGCACCATGTTCGAGGAGCGC	60	
Db	78562	ATGTGGGGCGCTTTTCTCTGCTCTGGGTCCGGGCGGACGCACCATGTTCGAGGAGCGC	78621	
QY	61	ACCATATTCGAGGACACCGCCCGCGGAGCGCGCGCAAGGACCCGCTGGGACACCTG	120	
Db	78622	ACCATATTCGAGGACACCGCCCGCGGAGCGCGCGCAAGGACCCGCTGGGACACCTG	78681	
QY	121	CGCACGGCGAGAAAGCGGGACCGTCCGGGGTGCTCCGGCGGCCCAACACCGTGTACCTG	180	
Db	78682	CGCACGGCGAGAAAGCGGGACCGTCCGGGGTGCTCCGGCGGCCCAACACCGTGTACCTG	78741	
QY	181	CAGTGTGTGGCAGCGGGTAGCGGGACTTCGGGCGCGCGGCTCTACGTCTTCTCCGAGTTC	240	
Db	78742	CAGTGTGTGGCAGCGGGTAGCGGGACTTCGGGCGCGCGGCTCTACGTCTTCTCCGAGTTC	78801	
QY	241	AACCGGTATCTCTTCAA	257	
Db	78802	AACCGGTAGTCAACGA	78818	

RESULT 14			
HSA329331	DNA	linear	PRI 18-JUL-2002
LOCUS	679 bp		
	HSA329331		

DEFINITION	Homo sapiens genomic sequence surrounding NotI site, clone NL6-B020C.									
ACCESSION	AJ329331									
VERSION	AJ329331.1									
KEYWORDS	GI:15873749									
SOURCE	Homo sapiens.									
ORGANISM	Homo sapiens									
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
AUTHORS	Kutsenko, A.S., Gizatullin, R.Z., Al-Amin, A.N., Wang, F., Kvasha, S.M., Podewski, R.M., Matushkin, Y.G., Gyanchandani, A., Muravenko, O.V., Levisky, V.G., Koshanov, N.A., Protopopov, A.I., Kashuba, V.I., Kisselev, L.L., Wasserman, W., Wahlstedt, C. and Zabarovsky, E.R.									
TITLE	NotI flanking sequences: a tool for gene discovery and verification of the human genome									
JOURNAL	Nucleic Acids Res. 30 (14), 3163-3170 (2002)									
PUBMED	12138098									
FEATURES	2 (bases 1 to 679)									
AUTHORS	Zabarovsky, E.R.									
TITLE	Direct Submission									
JOURNAL	Submitted (16-MAY-2001) Microbiology and Tumorbiology Centre, Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77, Sweden									
FEATURES	Location/Qualifiers									
source	1..679									
	/organism="Homo sapiens"									
	/db_xref="taxon:9606"									
	/clone="NL6-B020C"									
BASE COUNT	123 a 172 c 179 g 205 t									
ORIGIN										
Query Match	6.3%; Score 156.8; DB 9; Length 679;									
Best Local Similarity	95.8%; Pred. No. 1.1e-25;									
Matches 161; Conservative	0; Mismatches 7; Indels 0; Gaps 0;									
Qy	90 GCGGCCGCGCAAGACCCGCTGCGGCACCTGCGCACGCGAGAGAACGCGCGACGTCGGG 149									
Db	1 GCGGCCGCGCAAGACCCGCTGCGGCACCTGCGCACGCGAGAGAACGCGCGACGTCGGG 60									
Qy	150 GTGCTCGCGGGCGGCCAAACACCGTACCTGCGAGTGTGTGGCAGCGGAGTACCGGACTC 209									
Db	61 GTGCTCGCGGGCGGCCAAACACCGTACCTGCGAGTGTGTGGCAGCGGAGTACCGGACTC 120									
Qy	210 GGGCGCGCGCTCTACGCTCTTCCGAGTTCAACCGGTATCTCTTCAA 257									
Db	121 GGGCGCGCGCTCTACGCTCTTCCGAGTTCAACCGGTACGTAACGA 168									
RESULT 15										
AF304371S2										
LOCUS	AF304371S2 740 bp DNA linear PRI 19-OCT-2000									
DEFINITION	Homo sapiens putative prostate cancer susceptibility protein HPC2/ELAC2 gene, exon 17 and partial cds.									
ACCESSION	AF304369									
VERSION	AF304369.1									
KEYWORDS	GI:10880929									
SEGMENT	2 of 2									
SOURCE	Homo sapiens.									
ORGANISM	Homo sapiens									
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
AUTHORS	1 (bases 1 to 740)									
	Tavtigian, S.V., Simard, J., Teng, D.H.F., Baugard, M., Beck, A., Camp, N.J., Carillo, A.R., Chen, Y., Dayananth, P., Desrochers, M., Dumont, M., Farham, J.M., Frank, D., Frye, C., Ghaffari, S., Gupta, J.S., Hu, R., Iliev, D., Janacki, T., Kort, E.N., Lait, K.E., Leavitt, A., Leblanc, G., McArthur-Morrison, J., Pederson, A., Penn, B., Peterson, K.T., Reid, J.E., Richards, S., Schroeder, M., Smith, R., Snyder, S.C., Swedlund, B., Swensen, J., Thomas, A., Tranchant, M., Woodland, A.-M., Labrie, F., Skolnick, M.H., Neuhausen, S., Rommens, J. and Cannon-Albright, L.A.									
TITLE	A strong candidate prostate cancer susceptibility gene at									

RESULT 15	
AF304371S2	
LOCUS	740 bp DNA linear PRI 19-OCT-2000
DEFINITION	Homo sapiens putative prostate cancer susceptibility protein HPC2/ELAC2 gene, exon 17 and partial cds.
ACCESSION	AF304369
VERSION	AF304369.1 GI:10880929
KEYWORDS	.
SOURCE	2 of 2
ORGANISM	Homo sapiens.
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE	1 Tavtigian,S.V., Smard, <j., baumgard,m.,="" beck,a.,<br="" teng,d.h.f.,=""></j.,> Camp,N.J., Carillo,A.R., Chen,Y., Dayananath,P., Desrochers,M., Dumont,M., Farnham,J.M., Frank,D., Frye,C., Ghaffari,S., Gupte,J.S., Hu,R., Iliev,D., Janacki,T., Kort,E.N., Laity,K.E., Leavitt,A., Leblanc,G., McArthur-Morrison,J., Pederson,A., Penn,B., Peterson,K.T., Reid,J.E., Richards,S., Schroeder,M., Smith,R., Snyder,S.C., Swedlund,B., Swensen,J., Thomas,A., Tranchant,M., Woodland,A.-M., Labrie,F., Skolnick,M.H., Neuhausen,S., Rommens,J. and Cannon-Albright,L.A.
AUTHORS	A strong candidate prostate cancer susceptibility gene at
TITLE	

chromosome 17p
Unpublished
2 (bases 1 to 740)
JOURNAL
REFERENCE
AUTHORS
Tavtigian,S.V., Simard,J., Teng,D.H.F., Baumgard,M., Beck,A.,
Camp,N.J., Carillo,A.R., Chen,Y., Dayananth,P., Desrochers,M.,
Dumont,M., Farnham,J.M., Frank,D., Frye,C., Ghaffari,S.,
Gupte,J.S., Hu,R., Iliev,D., Janeski,T., Kort,E.N., Laity,K.E.,
Leavitt,A., Leblanc,G., McArthur-Morrison,J., Pederson,A., Penn,B.,
Peterson,K.T., Reid,J.E., Richards,S., Schroeder,M., Smith,R.,
Snyder,S.C., Swelund,B., Swensen,J., Thomas,A., Tranchant,M.,
Woodland,A.-M., Labrie,F., Skolnick,M.H., Neuhausen,S., Rommens,J.
and Cannon-Albright,L.A.
TITLE
JOURNAL
Direct Submission
Submitted (12-SEP-2000) Myriad Genetics, Inc., 320 Wakara Way, Salt
Lake City, UT 84108, USA
FEATURES
source
Location/Qualifiers
1..740
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17p"
mRNA
<301..>439
/product="putative prostate cancer susceptibility protein
HPC2/ELAC2"
CDS
<301..>439
/codon_start=2
/product="putative prostate cancer susceptibility protein
HPC2/ELAC2"
/protein_id="AAG24440.1"
/db_xref="GI:10880931"
/translation="PDTSLLLDCGEGTFGQLCRHYGQVDRVLGTLAAVFVSHLHADH
HT"
exon
301..439
/number=17
BASE COUNT 154 a 188 c 173 g 225 t
ORIGIN
Query Match 5.9%; Score 146.2; DB 9; Length 740;
Best Local Similarity 89.7%; Pred. No. 3.3e-23;
Matches 157; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
QY 1486 ATTGGAATGTGAGTCCACACTTGTCAACATAAGCCCGACACGCTCTGTCTACTGGAC 1545
DB 266 AGTGGCAGTGATCTCTTCCTCTTCTCTGACGCCCGACACGCTCTGTCTACTGGAC 325
QY 1546 TGTGTGAGGGCACATTTGGGCGAGCTGTGCCCTCATTTACGGAGACAGGTGACAGGGTTC 1605
DB 326 TGTGTGTGAGGGCACATTTGGGCGAGCTGTGCCCTCATTTACGGAGACAGGTGACAGGGTTC 385
QY 1606 CTGGGCACCTGGCTGCTGTGTTGTTGTTCCACCTGCAGCGAGATCACACACGG 1660
DB 386 CTGGGCACCTGGCTGCTGTGTTGTTGTTCCACCTGCAGCGAGATCACACACGG 440
Search completed: May 17, 2003, 21:25:41
Job time : 5600.12 secs

QY 241 AACCGGTATCTCTTCAACTGTGGAGAGCGCTTCAGAGACTCATGACAGGACACAAGTTA 300
DB |||||
QY 241 AACCGGTATCTCTTCAACTGTGGAGAGCGCTTCAGAGACTCATGACAGGACACAAGTTA 300
DB |||||
QY 301 AAGGTTGCTCGCCTGGACAACATATTCCTGACAGCAATGCACTGCTCTAAATGTTGGGGC 360
DB |||||
QY 301 AAGGTTGCTCGCCTGGACAACATATTCCTGACAGCAATGCACTGCTCTAAATGTTGGGGC 360
DB |||||
QY 361 TTAAGTGAATGATCTTTACTTTAAAGAAACCGGGCTTCCAAAGTGTAATCTTTCTGGA 420
DB |||||
QY 361 TTAAGTGAATGATCTTTACTTTAAAGAAACCGGGCTTCCAAAGTGTAATCTTTCTGGA 420
DB |||||
QY 421 CCTCCACAACCTGGAAAATACCTCGAAGCAATCAAAATATTTCTGTCATTTGAAGGA 480
DB |||||
QY 421 CCTCCACAACCTGGAAAATACCTCGAAGCAATCAAAATATTTCTGTCATTTGAAGGA 480
DB |||||
QY 481 ATAGAATCGGCTGTGCGGCCCACTCTCCCGCAGAAATACGAGGATGAAACCATGACAGTT 540
DB |||||
QY 481 ATAGAACTGGCTGTGCGGCCCACTCTCTCCCGCAGAAATACGAGGATGAAACCATGACAGTT 540
DB |||||
QY 541 TACAGATGCCATACACAGTGAAACAGAGGAGGAAAGCAACCAATGCGAGAGTCCA 600
DB |||||
QY 541 TACAGATGCCATACACAGTGAAACAGAGGAGGAAAGCAACCAATGCGAGAGTCCA 600
DB |||||
QY 601 GAAAGGCTCTCAGAGGCTCAGTCCAGAGCATCTTCAGACTCCGAGTCCGAATGAAAT 660
DB |||||
QY 601 GAAAGGCTCTCAGAGGCTCAGTCCAGAGCATCTTCAGACTCCGAGTCCGAATGAAAT 660
DB |||||
QY 661 GAGCCACACTTCCACATGGTGTAGCCAGAGAGAGGGGTGAGGACTCTTCCCTGGTC 720
DB |||||
QY 661 GAGCCACACTTCCACATGGTGTAGCCAGAGAGAGGGGTGAGGACTCTTCCCTGGTC 720
DB |||||
QY 721 GTAGCTTTATCTGTAGCTTCACTTAAAGAGAGAAACTTTTGGTGCTCAAAAGCAAG 780
DB |||||
QY 721 GTAGCTTTATCTGTAGCTTCACTTAAAGAGAGAAACTTTTGGTGCTCAAAAGCAAG 780
DB |||||
QY 781 GAGATGGGCTCCAGTTGGGACAGCTGCCATCGCTCCCATCATGCTGCTCAAGAC 840
DB |||||
QY 781 GAGATGGGCTCCAGTTGGGACAGCTGCCATCGCTCCCATCATGCTGCTCAAGAC 840
DB |||||
QY 841 GGGAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTCCA 900
DB |||||
QY 841 GGGAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTCCA 900
DB |||||
QY 901 GATCCTGTGTGCTTTTGTGGTGTAGAAATGTCAGATGAAAGCTTCATTCACCCATC 960
DB |||||
QY 901 GATCCTGTGTGCTTTTGTGGTGTAGAAATGTCAGATGAAAGCTTCATTCACCCATC 960
DB |||||
QY 961 TGTGAAATGCCACTTTTCAGAGGTACCAAGAAAGGACAGATGCCCGCTGGCTTGGTG 1020
DB |||||
QY 961 TGTGAAATGCCACTTTTCAGAGGTACCAAGAAAGGACAGATGCCCGCTGGCTTGGTG 1020
DB |||||
QY 1021 GTTCACATGGCCCAAGCATCTGTGCTGTGACAGAGGTACCAGAGTGGATGGAGAGG 1080
DB |||||
QY 1021 GTTCACATGGCCCAAGCATCTGTGCTGTGACAGAGGTACCAGAGTGGATGGAGAGG 1080
DB |||||
QY 1081 TTTGGGCTGACACCCAGCATTTGCTGCTGAATGAGAACTGTGCTCAGTTTCAACACTT 1140
DB |||||
QY 1081 TTTGGGCTGACACCCAGCATTTGCTGCTGAATGAGAACTGTGCTCAGTTTCAACACTT 1140
DB |||||
QY 1141 GCGAGCCACAGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCTTGGTC 1200
DB |||||
QY 1141 GCGAGCCACAGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCTTGGTC 1200
DB |||||
QY 1201 ACCAGTTTCCGCTGTAAAGAGAGGGGCCCACTTCACTGCTGCTGCTGCTGCTGCTGCT 1260
DB |||||
QY 1201 ACCAGTTTCCGCTGTAAAGAGAGGGGCCCACTTCACTGCTGCTGCTGCTGCTGCTGCT 1260
DB |||||
QY 1261 TGCCTCTCAAGTACCAAGTCCGTCAGGAGGGAGTGGCAGAGGGATGCCATTATTACT 1320
DB |||||
QY 1261 TGCCTCTCAAGTACCAAGTCCGTCAGGAGGGAGTGGCAGAGGGATGCCATTATTACT 1320
DB |||||
QY 1321 TGCATCTCTGAGGAATTCATAGTTGAGGGGCTGACGCTTCCCACTCCAGAGAGCGTG 1380
DB |||||

DB 1321 TGCATCTCTGAGGAATTCATAGTTGAGCGCTGCAGCTTCCCACTTCCAGCAGCGTG 1380
QY |||||
QY 1381 CAGAGTACAGGAGGAGTGCAGAGGACGGCCAGCCCCAGAGAGAGAGAGTCAAGTAC 1440
DB |||||
QY 1381 CAGAGTACAGGAGGAGTGCAGAGGACGGCCAGCCCCAGAGAGAGAGAGTCAAGTAC 1440
DB |||||
QY 1441 CCAGAAATCATCTTCTTGGAAACAGGCTTCCATCCCGATGAGATTCGAAATGTCACT 1500
DB |||||
QY 1441 CCAGAAATCATCTTCTTGGAAACAGGCTTCCATCCCGATGAGATTCGAAATGTCACT 1500
DB |||||
QY 1501 GCCACACTTGTCAACATAAGCCCCGACACAGTCTCTGCTACTTGGACTGTGGTGGGGCACA 1560
DB |||||
QY 1501 GCCACACTTGTCAACATAAGCCCCGACACAGTCTCTGCTACTTGGACTGTGGTGGGGCACA 1560
DB |||||
QY 1561 TTTGGGAGCTGTGCGCTCATTAACGAGACCAAGTGGACAGGGTCTTGGGACCCCTGGCT 1620
DB |||||
QY 1561 TTTGGGAGCTGTGCGCTCATTAACGAGACCAAGTGGACAGGGTCTTGGGACCCCTGGCT 1620
DB |||||
QY 1621 GCTGTGTTTGTGTCACCTGACAGGAGATCAACACAGGGCTTGGCAAGTATCTTGGTG 1680
DB |||||
QY 1621 GCTGTGTTTGTGTCACCTGACAGGAGATCAACACAGGGCTTGGCAAGTATCTTGGTG 1680
DB |||||
QY 1681 CAGAGAAACGCGCTTGGCATCTTTGGGAAAGCGCTTCAACCTTTGCTGGTGGTGGCC 1740
DB |||||
QY 1681 CAGAGAAACGCGCTTGGCATCTTTGGGAAAGCGCTTCAACCTTTGCTGGTGGTGGCC 1740
DB |||||
QY 1741 CCCAACAGCTCAAAGCTTGGCTCCAGCAGTACCAACACAGTGGCAGGAGGTCTCGAC 1800
DB |||||
QY 1741 CCCAACAGCTCAAAGCTTGGCTCCAGCAGTACCAACACAGTGGCAGGAGGTCTCGAC 1800
DB |||||
QY 1801 CACATCAGTATGATCTTCTGCAAAATGCCITTCAGGAAGGGCTGAGATCTCCAGTCTGCA 1860
DB |||||
QY 1801 CACATCAGTATGATCTTCTGCAAAATGCCITTCAGGAAGGGCTGAGATCTCCAGTCTGCA 1860
DB |||||
QY 1861 GTGGAAAGATTGATCAGTTCGCTGTTCCGAAACATGCTATTTGGAAGAGTTTTCAGACCTGT 1920
DB |||||
QY 1861 GTGGAAAGATTGATCAGTTCGCTGTTCCGAAACATGCTATTTGGAAGAGTTTTCAGACCTGT 1920
DB |||||
QY 1921 CTGCTGCGGCACTGCAAGCATGCTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
DB |||||
QY 1921 CTGCTGCGGCACTGCAAGCATGCTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
DB |||||
QY 1981 GTGCTCTATTCCGGGACACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
DB |||||
QY 1981 GTGCTCTATTCCGGGACACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
DB |||||
QY 2041 ACCCTCTGATACATGAAGCCACCTCGAAAGATGCTTTGGAAGAGGAGCAGTGGAAAG 2100
DB |||||
QY 2041 ACCCTCTGATACATGAAGCCACCTCGAAAGATGCTTTGGAAGAGGAGCAGTGGAAAG 2100
DB |||||
QY 2101 ACACAGCACAACGCTCCCAAGCCATCAGCGTGGGGATGCGGATGAAACGCGGAGTTCAAT 2160
DB |||||
QY 2101 ACACAGCACAACGCTCCCAAGCCATCAGCGTGGGGATGCGGATGAAACGCGGAGTTCAAT 2160
DB |||||
QY 2161 ATGCTGAACCATCTCAGCAGCGCTATGCCAAGTCCCTCTTCCAGCCCCCACTTCAGC 2220
DB |||||
QY 2161 ATGCTGAACCATCTCAGCAGCGCTATGCCAAGTCCCTCTTCCAGCCCCCACTTCAGC 2220
DB |||||
QY 2221 GAGAAAGTGGAGTTGCTTTGACCATGAAGTCTGCTTTGGAGACTTTTCCAAACATG 2280
DB |||||
QY 2221 GAGAAAGTGGAGTTGCTTTGACCATGAAGTCTGCTTTGGAGACTTTTCCAAACATG 2280
DB |||||
QY 2281 CCCAAGCTGATTTCCCACTGAAAGCCCTGTTGCTGCGGACATCGAGGAGATCGAGAG 2340
DB |||||
QY 2281 CCCAAGCTGATTTCCCACTGAAAGCCCTGTTGCTGCGGACATCGAGGAGATCGAGAG 2340
DB |||||
QY 2341 CGCAGGAGAGCGGGAGCTGCGGACAGTGGCGGCGGCTCTCTGCTCAGGGAGCTGGCA 2400
DB |||||
QY 2341 CGCAGGAGAGCGGGAGCTGCGGACAGTGGCGGCGGCTCTCTGCTCAGGGAGCTGGCA 2400
DB |||||
QY 2401 GCGGCTGAGGAGTGGGAGCTCAGCAGAGCGGCGCCACACAGAGGAGGCCACAGGCC 2460
DB |||||

Db	2401	GGCGCCTGGAGAGATGGGAGCCTCAGCAGAGAGCGGCCCA	CACAGAGGAGCCACAGGCC	2460
Qy	2461	AAGAAGGTCAGAGCCCCAGTGA	2481	
Db	2461	AAGAAGGTCAGAGCCCCAGTGA	2481	

RESULT 2

```

US-09-564-805-3
; Sequence 3, Application US/09564805
; Patent No. 633403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2958
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51)..(2531)
; OTHER INFORMATION: coding sequence as in SEQ ID NO:1
; US-09-564-805-3

```

	Query Match	100.0%;	Score 2481;	DB 4;	Length 2958;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 2481;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGTGGGCGCTTTGCTCGCTGCTGGCTTCGGGGCGGACGCACCATGTCCGAGGAGCGC	60		
Db	51	ATGTGGGCGCTTTGCTCGCTGCTGGCTTCGGGGCGGACGCACCATGTCCGAGGAGCGC	110		
Qy	61	ACCATATCGCAGGACACCGCCGCGCAGCGGGCGCGCAAGGACCCGCTGGCGCACCTG	120		
Db	111	ACCATATCGCAGGACACCGCCGCGCAGCGGGCGCGCAAGGACCCGCTGGCGCACCTG	170		
Qy	121	CGCAGCGAGAGAAGCGCGGACCGTTCGGGGTGCTCCGGCGGGCCCAAAACACCGTGTACCTG	180		
Db	171	CGCAGCGAGAGAAGCGCGGACCGTTCGGGGTGCTCCGGCGGGCCCAAAACACCGTGTACCTG	230		
Qy	181	CAGTGTGTGGCAGCGGGTAGCGGGACTCGGGCGCGCGCTCTACGTCTTCTCCAGTTC	240		
Db	231	CAGTGTGTGGCAGCGGGTAGCGGGACTCGGGCGCGCGCTCTACGTCTTCTCCAGTTC	290		
Qy	241	AACCGGTATCTTCAACTGTGGAGAAAGCGTTACAGACTCATCCAGGAGCACAAAGTTA	300		
Db	291	AACCGGTATCTTCAACTGTGGAGAAAGCGTTACAGACTCATCCAGGAGCACAAAGTTA	350		
Qy	301	AAGTTTGCTCGCCTGGACAAACATATTCCTGACACGAATGCACTGCTTAATGTTGGGGGC	360		
Db	351	AAGTTTGCTCGCCTGGACAAACATATTCCTGACACGAATGCACTGCTTAATGTTGGGGGC	410		
Qy	361	TTAAGTGAATGATCTTACTTTTAAGGAAACCGGGCTTCCAAAGTGCTACTTTCTCGGA	420		
Db	411	TTAAGTGAATGATCTTACTTTTAAGGAAACCGGGCTTCCAAAGTGCTACTTTCTCGGA	470		
Qy	421	CCTCCAACTGGAAAAATACCTCGAAGCAATCAAAAATATTTTCTGGTCCAATGAAGGA	480		
Db	471	CCTCCAACTGGAAAAATACCTCGAAGCAATCAAAAATATTTTCTGGTCCAATGAAGGA	530		

Qy	1561	TTTGGGACGCTGTGCCGTCA	TTTACGGAGACACAGGTGGA	ACGGGTCTCTGGGCA	CCCTGGCT	1620
Db	1611	TTTGGGACGCTGTGCCGTCA	TTTACGGAGACCAAGGTGGA	ACGGGTCTCTGGGCA	CCCTGGCT	1670
Qy	1621	GCTGTGTTTGTCGCCACCT	GTCAGCGCAGATACCA	CACGGGTCTGCCAAGTAT	CTTGCTG	1680
Db	1671	GCTGTGTTTGTCGCCACCT	GTCAGCGCAGATACCA	CACGGGTCTGCCAAGTAT	CTTGCTG	1730
Qy	1681	CAGAGAGAACCGCCCTTGG	CATCTTTGGGAAAGCCGCT	TACCCCTTCTCTGGTGGT	TGGCC	1740
Db	1731	CAGAGAGAACCGCCCTTGG	CACTCTTTGGGAAAGCCGCT	TACCCCTTCTCTGGTGGT	TGGCC	1790
Qy	1741	CCCAACGAGCTCAAGAGCT	TGGCTCAGCAGTACCA	CAACAGTGCAGAGGCT	CTCGAC	1800
Db	1791	CCCAACGAGCTCAAGAGCT	TGGCTCAGCAGTACCA	CAACAGTGCAGAGGCT	CTCGAC	1850
Qy	1801	CACATCAGTATGATCTCTCC	CAAAAGCCCTTACGAGGAG	GGGCTGAGATCTCCAGT	CCCTGCA	1860
Db	1851	CACATCAGTATGATCTCTCC	CAAAAGCCCTTACGAGGAG	GGGCTGAGATCTCCAGT	CCCTGCA	1910
Qy	1861	GTGGAAAGATTGATCAGTT	CGCTGTTTGGCAACATGT	GATTTGGAAGAGTTT	CAGACCTGT	1920
Db	1911	GTGGAAAGATTGATCAGTT	CGCTGTTTGGCAACATGT	GATTTGGAAGAGTTT	CAGACCTGT	1970
Qy	1921	CTGGTGGGCACTGGAAGCA	TGCTTGGCTGTGGCTG	GCACACCTCTGGCT	TGGAAA	1980
Db	1971	CTGGTGGGCACTGGAAGCA	TGCTTGGCTGTGGCTG	GCACACCTCTGGCT	TGGAAA	2030
Qy	1981	GTGGTCTATTTCGGGGACA	CCATGCCCTCGAGGCTCT	GTTCCGATGGGGAAGAT	GCC	2040
Db	2031	GTGGTCTATTTCGGGGACA	CCATGCCCTCGAGGCTCT	GTTCCGATGGGGAAGAT	GCC	2090
Qy	2041	ACCTCTCTGATACATGAAG	CCACCTCTGGAAGATGT	TTTGGAAAGAGACAGT	GCGAAAAG	2100
Db	2091	ACCTCTCTGATACATGAAG	CCACCTCTGGAAGATGT	TTTGGAAAGAGACAGT	GCGAAAAG	2150
Qy	2101	ACACACAGCAACAGTCCCA	AGCCATCAGCGTGGGATG	CGGATGAACCGGAGT	CTCAT	2160
Db	2151	ACACACAGCAACAGTCCCA	AGCCATCAGCGTGGGATG	CGGATGAACCGGAGT	CTCAT	2210
Qy	2161	ATGCTGAACACATTCAGC	CAGCGCTATGCCAAGGTC	CCCCCTCTTCAGCCCCA	CTTCAGC	2220
Db	2211	ATGCTGAACACATTCAGC	CAGCGCTATGCCAAGGTC	CCCCCTCTTCAGCCCCA	CTTCAGC	2270
Qy	2221	GAGAAAGTGGGAGTGTGCT	TTGACCAATGAAGGTC	TGCTTTGGAGACTTTT	CCAAACATG	2280
Db	2271	GAGAAAGTGGGAGTGTGCT	TTGACCAATGAAGGTC	TGCTTTGGAGACTTTT	CCAAACATG	2330
Qy	2281	CCCAAGCTGATTCCTCCAT	CTGAAGCCCTGTTTGT	TGTCGCGACATCGAG	GATGGAGG	2340
Db	2331	CCCAAGCTGATTCCTCCAT	CTGAAGCCCTGTTTGT	TGTCGCGACATCGAG	GATGGAGG	2390
Qy	2341	CGCAGGAGAACGGGAGCTG	CGGACAGTGGCGGGCC	CTCTCTCCAGGGAGCT	CGCA	2400
Db	2391	CGCAGGAGAACGGGAGCTG	CGGACAGTGGCGGGCC	CTCTCTCCAGGGAGCT	CGCA	2450
Qy	2401	GGCGGCTTGAGGATGGGAG	CCCTCAGCAGAAAGCGGG	CCCCCACAACAGAGAGCC	CAGGCC	2460
Db	2451	GGCGGCTTGAGGATGGGAG	CCCTCAGCAGAAAGCGGG	CCCCCACAACAGAGAGCC	CAGGCC	2510
Qy	2461	AAGAAGTCTAGAGCCAGTGA	2481			
Db	2511	AAGAAGTCTAGAGCCAGTGA	2531			

RESULT 3

RESUL 3
US-09-564-805-223
; Sequence 223, Application US/09564805
; Patent No. 633403
; GENERAL INFORMATION:
; APPLICANT: Tavtgian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques

Qy 721 GTAGCTTTTCATCTGTGAAGCTTCACTTAAAGAGAGAAACCTTTGGTGCTCAAGCAAG 780
Db 721 GTAGCTTTTCATCTGTGAAGCTTCACTTAAAGAGAGAAACCTTTGGTGCTCAAGCAAG 780
Qy 781 GAGATGGGCTCCAGTTGGGACAGCTGCCATCGTCCCATCATTTGCTGTCAAGGAC 840
Db 781 GAGATGGGCTCCAGTTGGGACAGCTGCCATCGTCCCATCATTTGCTGTCAAGGAC 840
Qy 841 GGGAAAGACATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTCCA 900
Db 841 GGGAAAGACATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTCCA 900
Qy 901 GATCCTGGTGTCTTTTGTGGTGGTGAATGTCCAGATGAAGCTTCATTCACCCCATC 960
Db 901 GATCCTGGTGTCTTTTGTGGTGGTGAATGTCCAGATGAAGCTTCATTCACCCCATC 960
Qy 961 TGTGAGATGACACCTTTTCAGAGTACCAAGAAAGGAGATGCCCGTGGCTTGGT 1020
Db 961 TGTGAGATGACACCTTTTCAGAGTACCAAGAAAGGAGATGCCCGTGGCTTGGT 1020
Qy 1021 GTTCACATGGCCCGACAGTCTGTGCTTGTGACAGCAGGTACACAGCTGGATGGAGGG 1080
Db 1021 GTTCACATGGCCCGACAGTCTGTGCTTGTGACAGCAGGTACACAGCTGGATGGAGGG 1080
Qy 1081 TTTGGGCTGACACCCAGCAGTCTGTCTCTGAATGAGAACTGTGCTCAGTTTCAACACTT 1140
Db 1081 TTTGGGCTGACACCCAGCAGTCTGTCTCTGAATGAGAACTGTGCTCAGTTTCAACACTT 1140
Qy 1141 CGAGCCCAAGATTCAAACCCAGCTCAACCTCATCCACCGGACATCTTCCCTGCTC 1200
Db 1141 CGAGCCCAAGATTCAAACCCAGCTCAACCTCATCCACCGGACATCTTCCCTGCTC 1200
Qy 1201 ACCAGTTTCCGCTGTAAGAAGAGGGCCCAACCTCAGTGTGCCATGTTTCAGGGTCAA 1260
Db 1201 ACCAGTTTCCGCTGTAAGAAGAGGGCCCAACCTCAGTGTGCCATGTTTCAGGGTCAA 1260
Qy 1261 TGCTCTCTCAAGTACCACTCCGTCAGGAGGGAGTGGCAGAGGATGCCATTATTACT 1320
Db 1261 TGCTCTCTCAAGTACCACTCCGTCAGGAGGGAGTGGCAGAGGATGCCATTATTACT 1320
Qy 1321 TGCAATCTGAGGAATTCATAGTTGAGGGCTGAGCTTCCCACTTCAGGAGAGCTG 1380
Db 1321 TGCAATCTGAGGAATTCATAGTTGAGGGCTGAGCTTCCCACTTCAGGAGAGCTG 1380
Qy 1381 CAGGATACAGGAGAGTGCAGCAGCAGCCAGCCAGCAGAGAGAAAGTCAAGTAC 1440
Db 1381 CAGGATACAGGAGAGTGCAGCAGCAGCCAGCCAGCAGAGAGAAAGTCAAGTAC 1440
Qy 1441 CCAGAAATCATCTTCTTTGGAACAGGGTCTGCCATCCCGATGGAAGATTCGAAATGTCAGT 1500
Db 1441 CCAGAAATCATCTTCTTTGGAACAGGGTCTGCCATCCCGATGGAAGATTCGAAATGTCAGT 1500
Qy 1501 GCCACACTTGTCAACATGAAGCCCGACAGCTCTCTGTACTGTGATCTGAGTGGAGGGACA 1560
Db 1501 GCCACACTTGTCAACATGAAGCCCGACAGCTCTCTGTACTGTGATCTGAGTGGAGGGACG 1560
Qy 1561 TTTGGGAGCTGTGCGCTCATTTACGAGACAGCTGAGCAGGGTCTCGGACACCTGGCT 1620
Db 1561 TTTGGGAGCTGTGCGCTCATTTACGAGACAGCTGAGCAGGGTCTCGGACACCTGGCT 1620
Qy 1621 GCTGTGTTTGTGTCCTCACTGACGACATCAACACAGGGCTTCCCAAGTATCTTGTG 1680
Db 1621 GCTGTGTTTGTGTCCTCACTGACGACATCAACACAGGGCTTCCCAAGTATCTTGTG 1680
Qy 1681 CAGAGAGAACCGCTTGTGCTTGTGGAAAGCCGCTTACCTTGTGCTGTGCTGTGCT 1740
Db 1681 CAGAGAGAACCGCTTGTGCTTGTGGAAAGCCGCTTACCTTGTGCTGTGCTGTGCT 1740
Qy 1741 CCCAACCCAGCTCAAGCTTGGCTCCAGCAGTACCAACACAGTGCAGAGGCTCTGCAC 1800
Db 1741 CCCAACCCAGCTCAAGCTTGGCTCCAGCAGTACCAACACAGTGCAGAGGCTCTGCAC 1800

Qy 1801 CACATCAGTATGATTTCTGCGCAATGCTTCAGGAAGGGCTGAGATCTCCAGTCTCGCA 1860
Db 1801 CACATCAGTATGATTTCTGCGCAATGCTTCAGGAAGGGCTGAGATCTCCAGTCTCGCA 1860
Qy 1861 GTGGAAGATTGATCAGTTCGCTGTTCGGAACATGTGATTTGGAAGAGTTTTCAGACCTGT 1920
Db 1861 GTGGAAGATTGATCAGTTCGCTGTTCGGAACATGTGATTTGGAAGAGTTTTCAGACCTGT 1920
Qy 1921 CTGGTGGGCACTGCAAGCATGCTTTGGCTGTGCTGTGTCACACTCTGTGGCTGGA 1980
Db 1921 CTGGTGGGCACTGCAAGCATGCTTTGGCTGTGCTGTGTCACACTCTGTGGCTGGA 1980
Qy 1981 GTGGTCTATTTCGGGGACACATGCCCTCGGAGCTCTGCTCCGATGGGAAAGATGCC 2040
Db 1981 GTGGTCTATTTCGGGGACACATGCCCTCGGAGCTCTGCTCCGATGGGAAAGATGCC 2040
Qy 2041 ACCCTCTGATACATGAAGCACCCTCGAAGATGTTTGAAGAGGAGCAGTGGAAAG 2100
Db 2041 ACCCTCTGATACATGAAGCACCCTCGAAGATGTTTGAAGAGGAGCAGTGGAAAG 2100
Qy 2101 ACACAGCACAACCTGCCAAGCCATCAGCGTGGGATGCGGATGAAACGCGGAGTTCA 2160
Db 2101 ACACAGCACAACCTGCCAAGCCATCAGCGTGGGATGCGGATGAAACGCGGAGTTCA 2160
Qy 2161 ATGCTGAACCACTTCAGCAGCGCTATGCCAAGTCCCTCTTCAGCCCCCACTTCAAC 2220
Db 2161 ATGCTGAACCACTTCAGCAGCGCTATGCCAAGTCCCTCTTCAGCCCCCACTTCAAC 2220
Qy 2221 GAGAAAGTGGGAGTTGCTTTGACCACATGAAGTCTGCTTTGGAGACTTTTCAACAATG 2280
Db 2221 GAGAAAGTGGGAGTTGCTTTGACCACATGAAGTCTGCTTTGGAGACTTTTCAACAATG 2280
Qy 2281 CCCAAGCTGATTTCCCACTGAAAGCCCTGTTTCTGCGCATCGAGGAGATGGAGAG 2340
Db 2281 CCCAAGCTGATTTCCCACTGAAAGCCCTGTTTCTGCGCATCGAGGAGATGGAGAG 2340
Qy 2341 CGCAGGAGAAAGCGGAGCTCGGCGAGTGGCGGGGCGCTCTCTGTCAGGAGCTGSCA 2400
Db 2341 CGCAGGAGAAAGCGGAGCTCGGCGAGTGGCGGGGCGCTCTCTGTCAGGAGCTGSCA 2400
Qy 2401 GCGGCTTGGAGATGGGAGCTCAGCAGAGCGGCGCCACACAGAGAGCCACAGGCC 2460
Db 2401 GCGGCTTGGAGATGGGAGCTCAGCAGAGCGGCGCCACACAGAGAGCCACAGGCC 2460
Qy 2461 AAGAAGGTTCAGAGCCCACTGA 2481
Db 2461 AAGAAGGTTCAGAGCCCACTGA 2481

RESULT 4

US-09-564-805-225
; Sequence 225, Application US/09564805
; Patent No. 633403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564.805
; PRIOR FILING DATE: 2000-05-05
; PRIOR FILING DATE: 1998-11-06
; PRIOR FILING DATE: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 225
; LENGTH: 2892
; TYPE: DNA

; ORGANISM: Gorilla gorilla

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(2478)

; US-09-564-805-225

Query Match 98.5%; Score 2442.6; DB 4; Length 2892;

Best Local Similarity 99.0%; Pred. No. 0;

Matches 2457; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy	1	ATGTGGGCGCTTTGCTCGCTGCTGCGGTCCGGCGCGGACGACCACTGTCGAGGAGCGC	60
Db	1	ATGTGGGCGCTTTGCTCGCTGCTGCGGTCCGGCGCGGACGACCACTGTCGAGGAGCGC	60
Qy	61	ACCATATCGCAGGACCCCGCGCGCGGAGCGCGCGCGCAAGGACCCGCTCGGGCACCTG	120
Db	61	ACCATATCGCAGGACCCCGCGCGCGGAGCGCGCGCGCAAGGACCCGCTCGGGCACCTG	120
Qy	121	CGCAGCGGAGAGAGCGCGGACCGTCCGGGTGCTCCGGCGCGCAACACACCGTGTACCTG	180
Db	121	CGCAGCGGAGAGAGCGCGGACCGTCCGGGTGCTCCGGCGCGCAACACACCGTGTACCTG	180
Qy	181	CAGGTGTGGCAGCGGGTAGCGGACTCGGGCGCGCGGCTCTACGTCTTCTCCGAGTTC	240
Db	181	CAGGTGTGGCAGCGGGTAGCGGACTCGGGCGCGCGGCTCTACGTCTTCTCCGAGTTC	240
Qy	241	AACCGGTATCTCTTCAACTGTGGAGAGCGGTTTCAAGACTCATGTCAGGAGCACAAAGTTA	300
Db	241	AACCGGTATCTCTTCAACTGTGGAGAGCGGTTTCAAGACTCATGTCAGGAGCACAAAGTTA	300
Qy	301	AAGGTGTGCTCGCTGGCAACATATTTCTGACAGAAATGCACTGCTTAATGTTGGGGC	360
Db	301	AAGGTGTGCTCGCTGGCAACATATTTCTGACAGAAATGCACTGCTTAATGTTGGGGC	360
Qy	361	TTAAGTGAATGATTTCTTACTTTTAAAGGAAACCGGGCTTCCAAAGTGTGACTTTCTGGA	420
Db	361	TTAAGTGAATGATTTCTTACTTTTAAAGGAAACCGGGCTTCCAAAGTGTGACTTTCTGGA	420
Qy	421	CCTCCAACTGGGAAATAATCTCGAAGAAATCAAAATATTTTCTGTCCTATGAAAGGA	480
Db	421	CCTCCAACTGGGAAATAATCTCGAAGAAATCAAAATATTTTCTGTCCTATGAAAGGA	480
Qy	481	ATAGAACTGGTGTGCGGCCCCACTCTCCCGAGAAATACGAGGATGAACCATGACAGTT	540
Db	481	ATAGAACTGGTGTGCGGCCCCACTCTCTCCCGAGAAATACGAGGATGAACCATGACAGTT	540
Qy	541	TACAGATCCCCATACACAGTGAACAGAGGAGGAGGAAAGCACCACCATGGCAGAGTCCA	600
Db	541	TACAGATCCCCATACACAGTGAACAGAGGAGGAGGAAAGCACCACCATGGCAGAGTCCA	600
Qy	601	GAAAGGCTCTCAGCAGGCTCAGTCCAGAGGATTTTCCAGACTCCGAGTCCGAATGAAAT	660
Db	601	GAAAGGCTCTCAGCAGGCTCAGTCCAGAGGATTTTCCAGACTCCGAGTCCGAATGAAAT	660
Qy	661	GAGCACAACCTTCCACATGTTAGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	720
Db	661	GAGCACAACCTTCCACATGTTAGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	720
Qy	721	GTAGCTTTTCACTGTAAAGCTTCACTTTAAAGAGAGGAAACTTCTTGGTCTCAAAGCAAG	780
Db	721	GTAGCTTTTCACTGTAAAGCTTCACTTTAAAGAGAGGAAACTTCTTGGTCTCAAAGCAAG	780
Qy	781	GAGATGGGCTCCCAAGTTGGGACAGCTGCCATCGCTCCCATCATTTGCTGTCAAGGAC	840
Db	781	GAGATGGGCTCCCAAGTTGGGACAGCTGCCATCGCTCCCATCATTTGCTGTCAAGGAC	840
Qy	841	GGGAAAGCATCATCATGAGGAGAGGATTTTGGCTGTGAGAGCTGTGTACTCTCTCA	900
Db	841	GGGAAAGCATCATCATGAGGAGAGGATTTTGGCTGTGAGAGCTGTGTACTCTCTCA	900
Qy	901	GATCCTGTGTCTTTTGTGTGTAGAAATGTCCAGATGAAAGTTCATTCACCCCATC	960
Db	901	GATCCTGTGTCTTTTGTGTGTAGAAATGTCCAGATGAAAGTTCATTCACCCCATC	960

Qy	961	TGTGAGAAATGCCACCTTTTACAGAGTACCAAGGAAAGCAGATGCCCGCTGGCCTTGGTG	1020
Db	961	TGTGAGAAATGCCACCTTTTACAGAGTACCAAGGAAAGCAGATGCCCGCTGGCCTTGGTG	1020
Qy	1021	GTTTACATGCGCCAGCATCTGTCTTGTGGACAGAGGTACCAAGAGTGGATGGAGAGG	1080
Db	1021	GTTTACATGCGCCAGCATCTGTCTTGTGGACAGAGGTACCAAGAGTGGATGGAGAGG	1080
Qy	1081	TTTGGGCTGACACCCAGCACTTGGTCTGAAATGAGAACTGTGCTCAGTTTCAAACTT	1140
Db	1081	TTTGGGCTGACACCCAGCACTTGGTCTGAAATGAGAACTGTGCTCAGTTTCAAACTT	1140
Qy	1141	CGGAGCCACAGATTCAAAACCCAGCTCAACTCATCCACCGGACATCTTCCCTGCTC	1200
Db	1141	CGGAGCCACAGATTCAAAACCCAGCTCAACTCATCCACCGGACATCTTCCCTGCTC	1200
Qy	1201	ACCAAGTTTCCGCTGTAAGAGAGGGGCCCAACCTCAGTGTGCCATGGTTTCAAGGTGAA	1260
Db	1201	ACCAAGTTTCCGCTGTAAGAGAGGGGCCCAACCTCAGTGTGCCATGGTTTCAAGGTGAA	1260
Qy	1261	TGCCTCTCAAGTACCAAGCTCCGTCCTCAGAGGAGGTGGCAGAGGATGCCATTATTACT	1320
Db	1261	TGCCTCTCAAGTACCAAGCTCCGTCCTCAGAGGAGGTGGCAGAGGATGCCATTATTACT	1320
Qy	1321	TGCAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCACTTCCAGCAGAGGTG	1380
Db	1321	TGCAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCACTTCCAGCAGAGGTG	1380
Qy	1381	CAGAGTACAGGAGGAGTGGCAGCGGCCAGCCAGCCAGCAGAGAAAGAAAGTCAAGTAC	1440
Db	1381	CAGAGTACAGGAGGAGTGGCAGCGGCCAGCCAGCCAGCAGAGAAAGAAAGTCAAGTAC	1440
Qy	1441	CCAGAAATCATCTTCTTGGAAACAGGCTCTGCCATCCCACTGAAAGTTCGAAATGTCACT	1500
Db	1441	CCAGAAATCATCTTCTTGGAAACAGGCTCTGCCATCCCACTGAAAGTTCGAAATGTCACT	1500
Qy	1501	GCCACACTTGTCAACATAAGCCCGGACAGCTCTCTGCTACTGGAATTCGAAATGTCACT	1560
Db	1501	GCCACACTTGTCAACATAAGCCCGGACAGCTCTCTGCTACTGGAATTCGAAATGTCACT	1560
Qy	1561	TTTGGGAGCTGTGCGCTCATACGAGAGACAGGTGACAGGCTCTGGGACCCCTGGCT	1620
Db	1561	TTTGGGAGCTGTGCGCTCATACGAGAGACAGGTGACAGGCTCTGGGACCCCTGGCT	1620
Qy	1621	GCTGTGTTGTGTCCTCCACTGACGAGATCAACACAGGCTTGCCTGCTGCTGCTGCTG	1680
Db	1621	GCTGTGTTGTGTCCTCCACTGACGAGATCAACACAGGCTTGCCTGCTGCTGCTGCTG	1680
Qy	1681	CAGAGAAACGCGCTTGGCATCTTTGGGAAAGCGCTTCACTTCTGCTGGTGGTGGCC	1740
Db	1681	CAGAGAAACGCGCTTGGCATCTTTGGGAAAGCGCTTCACTTCTGCTGGTGGTGGCC	1740
Qy	1741	CCCAACAGCTCAAAGCTGCTCCAGCAGTACCAACACAGTCCAGAGGCTCTGTCAC	1800
Db	1741	CCCAACAGCTCAAAGCTGCTCCAGCAGTACCAACACAGTCCAGAGGCTCTGTCAC	1800
Qy	1801	CACATCAGTATGATTCCTGCAAAATGCTTCCAGAGGAGGCTGAGATCTCCAGTCTGCA	1860
Db	1801	CACATCAGTATGATTCCTGCAAAATGCTTCCAGAGGAGGCTGAGATCTCCAGTCTGCA	1860
Qy	1861	GTGAAAGATGATCAGTTCGCTTTCGGAACATGATTTTGGAGAGTTCACACCTGT	1920
Db	1861	GTGAAAGATGATCAGTTCGCTTTCGGAACATGATTTTGGAGAGTTCACACCTGT	1920
Qy	1921	CTGTGCGGCACTCAAGCATGCGTTTGGCTGTGCTGTGACACCTCTGGCTGGAAA	1980
Db	1921	CTGTGCGGCACTCAAGCATGCGTTTGGCTGTGCTGTGACACCTCTGGCTGGAAA	1980
Qy	1981	GTGCTCTATTCCGGGACACCATGCCCTCGAGGCTCTGTGTCGGATGGGAAAGATGCC	2040
Db	1981	GTGCTCTATTCCGGGACACCATGCCCTCGAGGCTCTGTGTCGGATGGGAAAGATGCC	2040

Qy	2041	ACCTCTCTGATACATGAAGCCACCTCGNAGATGGTTTGGAGAGGAAACGCTGGAAAG	2100
Db	2041	ACCTCTCTGATACATGAAGCCACCTCGNAGATGGTTTGGAGAGGAAACGCTGGAAAG	2100
Qy	2101	ACACAGACACAACGTCCTCCAAAGCCATCAGCTGGGGATGCGGATGAACCGGGAGTTTCATT	2160
Db	2101	ACACAGACACAACGTCCTCCAAAGCCATCAGCTGGGGATGCGGATGAACCGGGAGTTTCATT	2160
Qy	2161	ATGCTGAACCACTTACGCAGCGCTATGCGAAGTCCCTCTTTCAGCCCAACTTCAGC	2220
Db	2161	ATGCTGAACCACTTACGCAGCGCTATGCGAAGTCCCTCTTTCAGCCCAACTTCAGC	2220
Qy	2221	GAGAAATGGGAGTTCCCTTTGACACACATGAAGTCTCTCTTTGGAGACTTTTCCAAACAATG	2280
Db	2221	GAGAAATGGGAGTTCCCTTTGACACACATGAAGTCTCTCTTTGGAGACTTTTCCAAACAATG	2280
Qy	2281	CCCAAGCTGATTTCCCCTCATGAAGCCCTGTTTGTGTCGCACATCGAGGAGATGGAGGAG	2340
Db	2281	CCCAAGCTGATTTCCCCTCATGAAGCCCTGTTTGTGTCGCACATCGAGGAGATGGAGGAG	2340
Qy	2341	CGCAGGAGNAGCGGAGCTGCGGCAGGTGCGGGCGGCCCTCTGTCCAGGGAGCTGGCA	2400
Db	2341	CGCAGGAGNAGCGGAGCTGCGGCAGGTGCGGGCGGCCCTCTGTCCAGGGAGCTGGCA	2400
Qy	2401	GGCGGCTCGAGGATGGGAGCCTCAGCAGAAAGCGGGCCCAACAGAGGAGCCACAGGCC	2460
Db	2401	GGCGGCTCGAGGATGGGAGCCTCAGCAGAAAGCGGGCCCAACAGAGGAGCCACAGGCC	2460
Qy	2461	AAGAGGTCAGAGCCCAAGTGA	2481
Db	2461	AAGAGGTCAGAGCCCAAGTGA	2481

```

RESULT 5
US-09-564-805-221
; Sequence 221, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtligian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Patalog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 221
; LENGTH: 2470
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2466)
US-09-564-805-221

```

Query Match	66.3%	Score 1645.6;	DB 4;	Length 2470;
Best Local Similarity	81.6%;	Pred. No. 0;		
Matches 1958;	Conservative	0;	Mismatches 417;	Indels 24; Gaps 4;

Qy	58	CGCACCATATTCGCAAGGACCCGCCGCCGAGCGGCCGCAAGACCGCTGCGGCAC	117
Db	40	CGCACCATGTGCAGGGTTCCGGCTCGTCGCCCGGCCACCCAAAGACCCACTGCGACAC	99
Qy	118	CTGCGCACGCGAGAGAACCGCGACCGTCGGGTGCTTCGCGCGGCCCAACACCGTGTAC	177

Db	100	CTCGGTACGGGGAGAAAGCGCGGCC-----GGGTCCGGGGGCCGGAACACCGTGTAC	153
Qy	178	CTCAGGTGTGTGGCAGCGGTTAGCGGGACTCGGGCGCGGCTCTACGTCTTCTCCGAG	237
Db	154	CTGCAGGTGTGGCGGCGGCGCGGAGCGCGGGGCTGCTCTCTATGTCTTCTCGGAA	213
Qy	238	TTCAA CGGTTATCTTTCAACTGTGGAGNAGGCGTTACAGACTCATGCAAGGACACAAG	297
Db	214	TACAAACAGGTACCTTTTAACTTCGGGAGAGCGCTCCAAACACTTATGCAAGAACACAAG	273
Qy	298	TTAAAGTTGCTCGCTCGACCAACATATTCCTGCACAGAAATGCACTGGTCTTAATGTTGGG	357
Db	274	ACTGAAGTCGCTCGCTTGACAACATCTTTCTGACTCGATGCGATTTGTTCAATGTTGGG	333
Qy	358	GGCTTAAGTGAATGATTTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTAATTTCT	417
Db	334	GGGTTGTGTGAATGATTTTAACTTTAAAGGAAACCGGGCTTCCAAATGTGTTCTGTCT	393
Qy	418	GGACCTCCACAACCTGGAATAATACCTCGNAGCAATCAAAATATTTTTCTGGTCCATTTGAA	477
Db	394	GGACCACCAACAGCTGAGAGAAATATCTAGAAGCAATCAAAATATTTTTCTGGTCCATTTGAA	453
Qy	478	GGATAGAACTGCTGTGCGGCCCACTCTGCCCCAGAAATACGAGGATGAAACCATGACA	537
Db	454	GGAAATAGAACTGCGCTGCGGCTCACTCTGCACAGAAATACAGGATGAGACCATGACT	513
Qy	538	GTTTACAGATCCCCATACACAGTGAACAGAGGGGAAAGCACCACCAATGGCAGAGT	597
Db	514	GTTTACAGGTCCCTATCCACAGTGAACGGAGTGTGGAAGCAACAGCCATCCAGAGC	573
Qy	598	CCAGAAAGGCTCTACAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCGAGTCSAATGAA	657
Db	574	CCAGAAACATCTCCCAACAGGCTCAGTCCCAACAGTCAATCGGACTCTGGATCAGCTGAA	633
Qy	658	AATGAGCCACCTTCCACATGGTGTAGCCAGAGAGGGGTCAAGGACTCTTCCCTG	717
Db	634	AATGGGC-----AGTGCACACAGGAAGCATGGGCGAGGAC-CCTCTCTTA	678
Qy	718	GTCTAGCTTTTCATCTGTAACTTCACTTAAAGAGAGAAACTTCTTGGTCTCAAAAGCA	777
Db	679	GTGGTAGCTTTTGTCTGCAAGCTTCACTTGAGGAAGAGAACTTCTTGGTCTTAAAGCA	738
Qy	778	AAGGAGATGGCTCCCGAGTGTGGACAGCTGCCATCGTCCCATCATTTGCTGTCTCAAG	837
Db	739	AAGGAGCTGGGCTTCTCTGTGGACGGCGGCAATGCACCCCATTCATTGCTGTCTCAAG	798
Qy	838	GACGGNAAAGCATCACTCATNAGGAAGAGAGATTTTGGCTGAGAGCTGTGTACTCCT	897
Db	799	GACGGNAAAGTATCACTTACGAAGGAAGAGAGATTCGTCTGAAGAGCTTTGTACACC	858
Qy	898	CCAGATCCTGGTCTCTTTTGTGGTGTAGAAATGTCAGATGAAAGCTTCATTCAACCC	957
Db	859	CCAGATCCTGGTCTTGTATTCACTGTGTGTAGTGTCTCTGATGAAGNATTCATCTCGCC	918
Qy	958	ATCTGTGAGAAATGCCACCTTTTACAGAGTACCAAGGAAAGGAGATGCCCGTGGCCCTTG	1017
Db	919	ATCTGTGAGAACGACACCTTTAAAGGTACAGGCGAGAGGCTGATGCACCTGTGGCGCTG	978
Qy	1018	GTGGTTCAATGCCCCCAGCATCTGTGCTTTGTGGACAGCAGGTACACAGCTGATGAG	1077
Db	979	GTGTTCCACATAGCCCCCAGAATCTGTACTCATCGACAGCAGATACCAAGCAGTGGATGG	1038
Qy	1078	AGTTTTGGGCTCACACCCAGCACTTGTCTCGAATGAGAACTGTGCCTCAGTTTCAACAC	1137
Db	1039	AGTTTCGGGCTTGACACACAGACCTGATTTCTGAATGAAATTTGCCCTTCGGTCCACAC	1098
Qy	1138	CTTCGACGCCACAAGATTCAAAACCCAGCTTCAACCTCATCCACCGGACATCTTCCCGCTG	1197
Db	1099	CTGCGACCCACAGATTACAGCCAGCTCAGCCTCATCCACCTTGACATCTTCCCCCAG	1158
Qy	1198	CTACACAGTTTTCGCTGTGAAGAGAGGGCCCCACCTCAGTGTGCCCATGTTGTCAGGGT	1257
Db	1159	CTTACACAGCTTCTATAGTAAGGAGGAAGGGTCCACCTTCAGCGTGCCTCAACAGTTTCGGGT	1218

Qy 1258 GAATGCTCTCAAGTACCAAGTCCGTCAGGAGGAGTGGCAGAGGATGCCATATT 1317
 Db 1219 GAATGCTCTCAAGTATTCAGTCCGCCCAAGAGAGAGTGGCAGAGGATACCAACATC 1278
 Qy 1318 ACTTGCATCTCGAGGAATTCATAGTTGAGGCGCTGCAAGCTTCCCACTCCAGCAGAGC 1377
 Db 1279 GACTGCATCTAGTGAATTCATAGTCAAGGCTTGGAGCTCCCAAGTTTCCAGAGAGT 1338
 Qy 1378 GTGAGGAGTACAGAGAGTGGCAGAGAGCGGCCAGCCCCAGCCAGCAGAGAGAGAGTCCAG 1437
 Db 1339 GTGAGGAGTATCGAAGAACCTGCAAGAAAACCCAGCCCCAGCAGAGAGAGAGAGCCAG 1398
 Qy 1438 TACCAGAAATCATCTCTTGGACAGGCTCGCATCCCGATCAAGATTCGAATGTC 1497
 Db 1399 TATCTGAAATTTCTCTCGGTACGGGCTCGCATCCCAATGAGATCCGAATGTC 1458
 Qy 1498 AGTGCCACACTTGTCAACATTAAGCCCCCAGACAGCTCTCTGCTACTGCACTGTGGTGAGGGC 1557
 Db 1459 AGTTCCACACTCGTCAACTAAGCCCTGACAAGTCAGTGTCTCTGATTTGGAGAGGC 1518
 Qy 1558 ACATTTGGCAGCTGTGCGCTCATTAACGAGACCAAGGTGGACAGGCTCTGGGCAACCTG 1617
 Db 1519 ACTTTGGCAGTGTGCGCTCAITACGACAGCAAAATAGACCGAGTCTTATGAGCCTC 1578
 Qy 1618 CTGCTGTGTTGTGTCCTCCACTGCAAGCAGATCACACAGGCTTCCCAAGTATCTTG 1677
 Db 1579 ACGGCTGTGTTGTGTCCTCCACTGCAAGCAGATCACACAGGCTTCCCAAGTATCTTG 1638
 Qy 1678 CTGAGAGAGAACGGCTTGGCATCTTTGGGAAAGCGCTTACCCCTTGTGCTGGTGT 1737
 Db 1639 CTGAGAGAGAGCATGCGTTGGCATCTCTGGGAAACCTTCCAGCCCTTGTGTTGGTG 1698
 Qy 1738 GCCCCCAACAGCTCAAGCTGGCTCCAGCAGTACCAACCAAGTGCAGAGAGTCCCTG 1797
 Db 1699 GCTCTACCCAGCTCAGGCGCTGCTGAGCAGTATCAACCACTGCAGAGATTCG 1758
 Qy 1798 CACCATCATGATGATTCCTGCCAAATCCCTTCCAGGAGGCTGAGATCTCCAGTCC 1857
 Db 1759 CACCATCATGATGATTCCTGCCAAATCCCTTCCAGGAGGCTGAGATCTCCAAATCT 1818
 Qy 1858 GCAGTGGAAAGATTCATGATGCTGCTTGGCAACATGTGATTTGGAAGAGTTTCAGACC 1917
 Db 1819 ACATTTGGAAGGCTGATAGCTTGTGTTGGAAACATGTGATTTAGAGAAATTCAGACC 1878
 Qy 1918 TGTCTGGTGGGCACTGCAAGCATGCTGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1977
 Db 1879 TGCCTGGTACGGCACTGCAAGCATGCTTTGGCTGTGCACTGCTGCTGCTGCTGCTGCTG 1938
 Qy 1978 AAAGTGTCTATTCCGGGACACCATGCCCTCGAGGCTCTGCTCGGATGGGAAAGAT 2037
 Db 1939 AAAGTGTCTATTCCGGGAGTACCATGCCCTGTGAGGCTCTGCTCGGATGGGAAAGAT 1998
 Qy 2038 GCCACCTCTGATACATGAAGCCACCTCGAAGATGTTTGGAAAGAGAGCAGTGGAA 2097
 Db 1999 GCCACCTCTGATACATGAAGCCACTCTGAGGATCNCCTTGGAAAGAGAGCAGTAGAG 2058
 Qy 2098 AAGACACAGACCAACGTGCCAAGCATACAGCTGGGGATGCGATGAACCGGAGTTC 2157
 Db 2059 AGGACACACAGACCAACCTCCAGGCTATTAATGTGGGGATCGGATGAATCGGAGTTC 2118
 Qy 2158 ATTATGCTGAACCACTTCAGCCAGGCTATGCCAAGTCCCTCTTCCAGCCCACTTC 2217
 Db 2119 ATCATGTGAACCACTTCAGTCAAGGCTACGCNAAAGATCCCTCTTTCAGCCCTGACTTC 2178
 Qy 2218 AGCGAGAAAGTGGAGTTCCTTTGACACATGAAGGCTGCTTTGGAGACTTTTCCAAACA 2277
 Db 2179 AACGAGAAAGTGGCATCGCTTTGACACATGAAGTCTGNTTTGGAGACTTCCGACA 2238
 Qy 2278 ATGCCCAAGCTGATTTCCCACTGAAGCCCTGTTGCTGGCGCATCATGAGAGATGAG 2337
 Db 2239 GTGCCCAAGCTGATTTCCCACTGAAGCCCTGTTTGCAGGAGTGCATTTGAAGAGATGGTG 2298

Qy 2338 GAGCGAGGAGAGAGCGGAGCTCGGCGAGTGGCGGCGCCCTCTCTGTCCAGGAGCTG 2397
 Db 2299 GAACGAGGAGAGAGAGGAGCTACGGCTGGTGGAGCAGCCCTCTCTGACC---CAGCAG 2355
 Qy 2398 GCAGGCGCTTGGAGGATGGGAGCTTACAGAGAGCGGCGCCACACAGAGGAGCCACA 2456
 Db 2356 GCAGACGCGCAGAGGACAGAGAACCCCAACAGAGCGGCGCCACACAGATGAACACACA 2414

RESULT 6
 US-09-564-805-210
 ; Sequence 210, Application US/09564805
 ; Patent No. 6333403
 ; GENERAL INFORMATION:
 ; APPLICANT: Tavtigian, Sean V.
 ; APPLICANT: Teng, David H.F.
 ; APPLICANT: Simard, Jacques
 ; APPLICANT: Rommens, Johanna M.
 ; APPLICANT: Myriad Genetics, Inc.
 ; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
 ; FILE REFERENCE: Gene and a Paralog and Orthologous Genes
 ; CURRENT APPLICATION NUMBER: US/09/564,805
 ; CURRENT FILING DATE: 2000-05-05
 ; PRIOR APPLICATION NUMBER: US 60/107,468
 ; PRIOR FILING DATE: 1998-11-06
 ; PRIOR APPLICATION NUMBER: 09/434,382
 ; PRIOR FILING DATE: 1999-11-05
 ; NUMBER OF SEQ ID NOS: 240
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 210
 ; LENGTH: 350
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (51)..(293)
 US-09-564-805-210

Query Match 10.0%; Score 247.4; DB 4; Length 350;
 Best Local Similarity 97.7%; Pred. No. 2.1e-57;
 Matches 251; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 ATGTGGCGCTTCTCGCTCGCTGCTCGCGCGCGGAGCAGCAGCAGCAGTGTCCAGGAGCGC 60
 Db 51 ATGTGGCGCTTCTCGCTCGCTGCTCGCGCGCGGAGCAGCAGCAGCAGTGTCCAGGAGCGC 110
 Qy 61 ACCATATCGCAGGACCAACCGCGCGCGGAGCGCGCGGAGCAGCAGCAGCAGCAGTGTCCGCGACCTG 120
 Db 111 ACCATATCGCAGGACCAACCGCGCGCGGAGCGCGCGGAGCAGCAGCAGCAGTGTCCGCGACCTG 170
 Qy 121 CGCAGCGGAGAGAGCGCGGAGCGCGCGGAGCGCGCGGAGCAGCAGCAGCAGTGTCCGCGACCTG 180
 Db 171 CGCAGCGGAGAGAGCGCGGAGCGCGCGGAGCGCGCGGAGCAGCAGCAGCAGTGTCCGCGACCTG 230
 Qy 181 CAGCTGTGGCAGCGGCTAGCCGGAGCTCGGGCGCGCGGAGCAGCAGCAGCAGTGTCCGAGTTC 240
 Db 231 CAGCTGTGGCAGCGGCTAGCCGGAGCTCGGGCGCGCGGAGCAGCAGCAGCAGTGTCCGAGTTC 290
 Qy 241 AACCGGTATCTCTTCAA 257
 Db 291 AACCGGTATCTCTTCAA 307

RESULT 7
 US-09-564-805-28
 ; Sequence 28, Application US/09564805
 ; Patent No. 6333403
 ; GENERAL INFORMATION:
 ; APPLICANT: Tavtigian, Sean V.
 ; APPLICANT: Teng, David H.F.
 ; APPLICANT: Simard, Jacques
 ; APPLICANT: Rommens, Johanna M.

RESULT 8
US-09-564-805-4
; Sequence 4, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:

/ PATENT NO.: 6202353
 / GENERAL INFORMATION:
 / APPLICANT: Endege, Wilson O.
 / APPLICANT: Steinmann, Kathleen E.
 / APPLICANT: Astle, Jon H.
 / APPLICANT: Burgess, Christopher C.
 / APPLICANT: Bushnell, Steven E.
 / APPLICANT: Carroll III, Eddie
 / APPLICANT: Catino, Theodore J.
 / APPLICANT: Derti, Adnan
 / APPLICANT: Ford, Donna M.
 / APPLICANT: Lewis, Marcia E.
 / APPLICANT: Monahan, John E.
 / APPLICANT: Schlegel, Robert
 / TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
 / TITLE OF INVENTION: PRODUCTS
 / FILE REFERENCE: CCD-257 (US)
 / CURRENT APPLICATION NUMBER: US/09/328,111
 / CURRENT FILING DATE: 1999-06-08

; EARLIER APPLICATION NUMBER: US 60/088,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 315
; LENGTH: 238
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-328-111-315

Query Match 9.6%; Score 237; DB 4; Length 238;
Best Local Similarity 100.0%; Pred. No. 1.le-54;
Matches 237; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 176 ACCTGCAGGTGGTGCACGGGTAGCCGGGACTCGGGCGCGCGCTCTACGTCTTCTCCG 235
Db 1 ACCTGCAGGTGGTGCACGGGTAGCCGGGACTCGGGCGCGCGCTCTACGTCTTCTCCG 60
Qy 236 AGTTCAACCGGTATCTTCAACTGTGAGAAGGGGTTTCAGAGACTCATGCGAGGACACA 295
Db 61 AGTTCAACCGGTATCTTCAACTGTGAGAAGGGGTTTCAGAGACTCATGCGAGGACACA 120
Qy 296 AGTTAAAGTTGCTCGCTCGCCTGGACACATATTCCTGACACGAATGCATGGTCTAATGTTG 355
Db 121 AGTTAAAGTTGCTCGCTCGCCTGGACACATATTCCTGACACGAATGCATGGTCTAATGTTG 180
Qy 356 GGGGCTTAAGTGGATGATCTTACTTTAAAGGAACCGGGCTTCCAAAGTGTGTAC 412
Db 181 GGGGCTTAAGTGGATGATCTTACTTTAAAGGAACCGGGCTTCCAAAGTGTGTAC 237

RESULT 10

US-09-564-805-27
; Sequence 27, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 655
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(228)
; OTHER INFORMATION: exon 24
; NAME/KEY: polyA signal
; LOCATION: (636)..(641)
US-09-564-805-27

Query Match 9.2%; Score 228; DB 4; Length 655;
Best Local Similarity 100.0%; Pred. No. 5.le-52;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2254 GTCCTCTTGGAGCTTCCCAATGCCAGCTGATCCCCCACTGAAGCCCTGTTT 2313
Db 1 GTCCTCTTGGAGCTTCCCAATGCCAGCTGATCCCCCACTGAAGCCCTGTTT 60
Qy 2314 GCTGCGACATCGAGGATGAGGAGCGCAGGAGAAAGCTCGGCGAGGTGCGG 2373

Db 61 GCTGCGACATCGAGGAGATGGAGGCGCAGGAGAAGCGGAGCTGCGCAGGTGCGG 120
Qy 2374 CGGCGCTCTCTGTCAGGAGCTGGCAGGCGGCTCGAGGATGGGAGCCTCAGCAGAAG 2433
Db 121 CGGCGCTCTCTGTCAGGAGCTGGCAGGCGGCTCGAGGATGGGAGCCTCAGCAGAAG 180
Qy 2434 CGGCGCCACACAGAGGAGCCACAGGGCCAAAGGTTCAGAGCCCAAGTGA 2481
Db 181 CGGCGCCACACAGAGGAGCCACAGGGCCAAAGGTTCAGAGCCCAAGTGA 228

RESULT 11

US-09-564-805-26
; Sequence 26, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
; LENGTH: 145
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(145)
; OTHER INFORMATION: exon 23
US-09-564-805-26

Query Match 5.8%; Score 145; DB 4; Length 145;
Best Local Similarity 100.0%; Pred. No. 6e-30;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2109 CACAACGTCCCAAGCCATCAGCGTGGGGATGCGGATGAACGCGGAGTTCAATTATGCTGAA 2168
Db 1 CACAACGTCCCAAGCCATCAGCGTGGGGATGCGGATGAACGCGGAGTTCAATTATGCTGAA 60
Qy 2169 CCATTTCAGCCAGCGCTATGCCAAGTCCCGCTCTTCAGCCCCCACTTCACGCGAAGT 2228
Db 61 CCATTTCAGCCAGCGCTATGCCAAGTCCCGCTCTTCAGCCCCCACTTCACGCGAAGT 120
Qy 2229 GGGAGTTGCCCTTTGACCACATGAAG 2253
Db 121 GGGAGTTGCCCTTTGACCACATGAAG 145

RESULT 12

US-09-564-805-16
; Sequence 16, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258

; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 139
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(139)
; OTHER INFORMATION: exon 13
US-09-564-805-16

Query Match 5.6%; Score 139; DB 4; Length 139;
Best Local Similarity 100.0%; Pred. No. 2.5e-28;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1080 GTTTGGGCTGACACCCAGCACTTGGTCTGTAATGAGAACTGTGCTCAGTTCAACAACCT 1139
Db 1 GTTTGGGCTGACACCCAGCACTTGGTCTGTAATGAGAACTGTGCTCAGTTCAACAACCT 60
Qy 1140 TCGCAGCCACAAAGATTCAAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCCCTGCT 1199
Db 61 TCGCAGCCACAAAGATTCAAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCCCTGCT 120
Qy 1200 CACCAAGTTCCGCTGAAG 1218
Db 121 CACCAAGTTCCGCTGAAG 139

RESULT 13
US-09-564-805-20
; Sequence 20, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 139
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(139)
; OTHER INFORMATION: exon 17
US-09-564-805-20

Query Match 5.6%; Score 139; DB 4; Length 139;
Best Local Similarity 100.0%; Pred. No. 2.5e-28;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1521 CCCCACAGCTCTGCTACTGAGTGGTGGGACCAATTTGGGCAGCTGTGCGGTCA 1580
Db 1 CCCCACAGCTCTGCTACTGAGTGGTGGGACCAATTTGGGCAGCTGTGCGGTCA 60

Qy 1581 TTACGGAGACCAGGTGGACAGGGTCCTGGGCAACCTGCTGCTGTGTGTGTCACCACT 1640
Db 61 TTACGGAGACCAGGTGGACAGGGTCCTGGGCAACCTGCTGCTGTGTGTGTCACCACT 120
Qy 1641 GCAGCGAGATCACCACACG 1659
Db 121 GCAGCGAGATCACCACACG 139

RESULT 14
US-09-564-805-24
; Sequence 24, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 121
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(121)
; OTHER INFORMATION: exon 21
US-09-564-805-24

Query Match 4.9%; Score 121; DB 4; Length 121;
Best Local Similarity 100.0%; Pred. No. 1.7e-23;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1909 TTTCAGACCTCTGTGTCGGCAGCTGCAAGCATGCGTTTGGCTGTGCGCTGTGTCACACC 1968
Db 1 TTTCAGACCTCTGTGTCGGCAGCTGCAAGCATGCGTTTGGCTGTGCGCTGTGTCACACC 60

Qy 1969 TCTGGCTGGAAGTGGTCTATTCCGGGGACACCATGCCCTCGAGGCTCTGTCGGGATG 2028
Db 61 TCTGGCTGGAAGTGGTCTATTCCGGGGACACCATGCCCTCGAGGCTCTGTCGGGATG 120
Qy 2029 G 2029
Db 121 G 121

RESULT 15
US-09-564-805-10
; Sequence 10, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05

```
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver..2.0
; SEQ ID NO 10
; LENGTH: 120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(120)
; OTHER INFORMATION: exon 7
US-09-564-805-10

Query Match      4.8%; Score 120; DB 4; Length 120;
Best Local Similarity 100.0%; Pred.No. 3.le-23;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 560 GTGAACAGAGGAGGGAAGCACCACCAACCATGGCAGAGTCCAGAAAGGCGCTCTCAGCAGGC 619
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1  GTGAACAGAGGAGGGAAGCACCACCAACCATGGCAGAGTCCAGAAAGGCGCTCTCAGCAGGC 60

Qy 620 TCAGTCCAGAGCGATCTTCAGACTCCGAGTCGAATGAAAATGAGCCACACCTTCCACATG 679
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 TCAGTCCAGAGCGATCTTCAGACTCCGAGTCGAATGAAAATGAGCCACACCTTCCACATG 120
```

Search completed: May 17, 2003, 16:40:30
Job time : 153.603 secs

Result No.	Query %			DB	ID	Description
	Score	Match	Length			
1	2481	100.0	2481	9	US-09-988-626-1	Sequence 1, Appli
2	2481	100.0	2481	9	US-09-988-687-1	Sequence 1, Appli
3	2481	100.0	2958	9	US-09-988-626-3	Sequence 3, Appli
4	2481	100.0	2958	9	US-09-988-687-3	Sequence 3, Appli
5	2455.4	99.0	2908	9	US-09-988-626-223	Sequence 223, App
6	2455.4	99.0	2908	9	US-09-988-687-223	Sequence 223, App
7	2442.6	98.5	2892	9	US-09-988-626-225	Sequence 225, App
8	2442.6	98.5	2892	9	US-09-988-687-225	Sequence 225, App
9	1445.6	66.3	2470	9	US-09-988-626-221	Sequence 221, App
10	1445.6	66.3	2470	9	US-09-988-687-221	Sequence 221, App
11	734.8	29.6	783	10	US-09-833-381-2039	Sequence 2039, A
12	470.4	19.0	536	10	US-09-833-381-2038	Sequence 2038, A
13	432.8	17.4	554	9	US-09-918-995-8996	Sequence 8996, Ap
14	247.4	10.0	350	9	US-09-988-626-210	Sequence 210, App
15	247.4	10.0	350	9	US-09-988-687-210	Sequence 210, App
16	247.4	10.0	26684	9	US-09-988-626-28	Sequence 28, Appli
17	247.4	10.0	26684	9	US-09-988-687-28	Sequence 28, Appli
18	245	9.9	295	9	US-09-988-626-4	Sequence 4, Appli
19	245	9.9	295	9	US-09-988-687-4	Sequence 4, Appli

Db 61 ACCATATCGAGGACCCGCGCGCGAGCGCGCGCAAGGACCCGCTGCGGCACTG 120
Qy 121 CGCAGCGAGAGAACGGGACCGTTCGGGTGCTCCGGGGGCCCAAAACACGGTGTACCTG 180
Db 121 CGCAGCGAGAGAACGGGACCGTTCGGGTGCTCCGGGGGCCCAAAACACGGTGTACCTG 180
Qy 181 CAGGTGTGGCAGCGGGTAGCCGGACTCGGGCGCGCGCTCTACGTCTTCTCCGAGTTC 240
Db 181 CAGGTGTGGCAGCGGGTAGCCGGACTCGGGCGCGCGCTCTACGTCTTCTCCGAGTTC 240
Qy 241 AACCGGTATCTCTTCAACTGTGGAGAGCGTTTCAGAGACTCATCAGGAGCACAAAGTTA 300
Db 241 AACCGGTATCTCTTCAACTGTGGAGAGCGTTTCAGAGACTCATCAGGAGCACAAAGTTA 300
Qy 301 AAGTGTGCTCGCTGGACAAATATCTCTGACACGAATGCACTGTCTAAATTTGGGGGC 360
Db 301 AAGTGTGCTCGCTGGACAAATATCTCTGACACGAATGCACTGTCTAAATTTGGGGGC 360
Qy 361 TTAAGTGAATGATCTTACTTTTAAAGGAAACCGGGCTTCCAAAGTGTAATTTCTGGA 420
Db 361 TTAAGTGAATGATCTTACTTTTAAAGGAAACCGGGCTTCCAAAGTGTAATTTCTGGA 420
Qy 421 COTCCAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTTGAAAGGA 480
Db 421 COTCCAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTTGAAAGGA 480
Qy 481 ATAGAACTGGTGTGCGGCCCACTCTGCGCCAGAAATACGAGGATGAAACCATGACATT 540
Db 481 ATAGAACTGGTGTGCGGCCCACTCTGCGCCAGAAATACGAGGATGAAACCATGACATT 540
Qy 541 TACAGATCCCATACAGATGAACAGAGAGGGGAAAGCAACCACTGACAGTCCA 600
Db 541 TACAGATCCCATACAGATGAACAGAGAGGGGAAAGCAACCACTGACAGTCCA 600
Qy 601 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCGAATGAAAT 660
Db 601 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCGAATGAAAT 660
Qy 661 GAGCCACACTTCCACATGTGTTAGCCAGAGAGAGGGGTACAGGAATCTTCCCTGTGTC 720
Db 661 GAGCCACACTTCCACATGTGTTAGCCAGAGAGAGGGGTACAGGAATCTTCCCTGTGTC 720
Qy 721 GTAGCTTTTCACTGTAAGCTTCACTTAAAGAGAGAAACTTCTTGGTGTCAAAGCAAG 780
Db 721 GTAGCTTTTCACTGTAAGCTTCACTTAAAGAGAGAAACTTCTTGGTGTCAAAGCAAG 780
Qy 781 GAGATGGGCTCCCACTGGGACAGCTGCCATCGCTCCCATCAATGCTGTCAAAGAC 840
Db 781 GAGATGGGCTCCCACTGGGACAGCTGCCATCGCTCCCATCAATGCTGTCAAAGAC 840
Qy 841 GGGAAAGCATCATCATGAAGGAAGAGATTTTGGCTGAAGAGCTGTGTACTCTCCA 900
Db 841 GGGAAAGCATCATCATGAAGGAAGAGATTTTGGCTGAAGAGCTGTGTACTCTCCA 900
Qy 901 GATCCTGTGTCTTTTGTGTGTAGAACTGTCAGATGAAGCTTCAATCAACCCATC 960
Db 901 GATCCTGTGTCTTTTGTGTGTAGAACTGTCAGATGAAGCTTCAATCAACCCATC 960
Qy 961 TGTGAGAAATGCCACTTTTCAAGGTACCAAGGAAAGGAGATGCCCGCTGGCTTGGTG 1020
Db 961 TGTGAGAAATGCCACTTTTCAAGGTACCAAGGAAAGGAGATGCCCGCTGGCTTGGTG 1020
Qy 1021 GTTTCATGTGGCCCAAGCATCTGTGTTGTGGACAGAGGTACACAGTGTGATGGAGG 1080
Db 1021 GTTTCATGTGGCCCAAGCATCTGTGTTGTGGACAGAGGTACACAGTGTGATGGAGG 1080
Qy 1081 TTTGGGCTGACACCCAGCATTTGTCTCTGAATGAACTGTGCTCAGTTTCAACACTT 1140
Db 1081 TTTGGGCTGACACCCAGCATTTGTCTCTGAATGAACTGTGCTCAGTTTCAACACTT 1140
Qy 1141 CGCAGCCAAAGATTCAAAACCCAGCTCAACCTTCATCCACCCGGACATCTCCCTGTC 1200
Db 1141 CGCAGCCAAAGATTCAAAACCCAGCTCAACCTTCATCCACCCGGACATCTTCCCTGTC 1200

Qy 1201 ACCAGTTTCGCTGTAAAGAGGGGCCCCACCTCAGTGTGCCCATGGTTTCAGGTTGAA 1260
Db 1201 ACCAGTTTCGCTGTAAAGAGGGGCCCCACCTCAGTGTGCCCATGGTTTCAGGTTGAA 1260
Qy 1261 TGCCTCTCAAGTACCAGTCCGTCAGAGGAGGTGGCAGAGGATGCCATTATTACT 1320
Db 1261 TGCCTCTCAAGTACCAGTCCGTCAGAGGAGGTGGCAGAGGATGCCATTATTACT 1320
Qy 1321 TGCATCTGAGGAATTCATAGTTGAGGCGTGCAGCTTCCCACTTCCAGCAGAGCGTG 1380
Db 1321 TGCATCTGAGGAATTCATAGTTGAGGCGTGCAGCTTCCCACTTCCAGCAGAGCGTG 1380
Qy 1381 CAGAGTACAGGAGAGTGGCAGGACGGCCAGCCAGCCAGCAGAGAGAAAGTCAGTAC 1440
Db 1381 CAGAGTACAGGAGAGTGGCAGGACGGCCAGCCAGCCAGCAGAGAGAAAGTCAGTAC 1440
Qy 1441 CCAGAAATCATCTTCTTGGAAACAGGGTCTGCCATCCGATGAAAGATTTCGAAATGTCA 1500
Db 1441 CCAGAAATCATCTTCTTGGAAACAGGGTCTGCCATCCGATGAAAGATTTCGAAATGTCA 1500
Qy 1501 GCCACACTTGTCAACATAAGCCCCGACACGCTCTCTGCTACTTGGACTGTGTGAGGGCACA 1560
Db 1501 GCCACACTTGTCAACATAAGCCCCGACACGCTCTCTGCTACTTGGACTGTGTGAGGGCACA 1560
Qy 1561 TTTGGGAGCTGTGCGCTCATACGGAGACAGGTGGACAGGGTCTCTGGGCACTTGGCT 1620
Db 1561 TTTGGGAGCTGTGCGCTCATACGGAGACAGGTGGACAGGGTCTCTGGGCACTTGGCT 1620
Qy 1621 GCTGTGTTGTGTCCTCCACCTGACGAGATCAACACAGGCTTCCCAAGTATCTTGTG 1680
Db 1621 GCTGTGTTGTGTCCTCCACCTGACGAGATCAACACAGGCTTCCCAAGTATCTTGTG 1680
Qy 1681 CAGAGAAACGCGCTTGGCATCTTTGGGAAAGCGCTTCACTTCTGCTGTTGCTG 1740
Db 1681 CAGAGAAACGCGCTTGGCATCTTTGGGAAAGCGCTTCACTTCTGCTGTTGCTG 1740
Qy 1741 CCCAACAGCTCAAAAGCTGGCTCCAGCAGTACCAACACAGTCCAGGAGGTCTGAC 1800
Db 1741 CCCAACAGCTCAAAAGCTGGCTCCAGCAGTACCAACACAGTCCAGGAGGTCTGAC 1800
Qy 1801 CACATCAGTATGATTCTGTCCTCAATGCTTTCAGGAGGGCTGAGATCTCCAGTCTGCA 1860
Db 1801 CACATCAGTATGATTCTGTCCTCAATGCTTTCAGGAGGGCTGAGATCTCCAGTCTGCA 1860
Qy 1861 GTGAAAGATTGATCAGTTCGCTGTTCGAAACATGTGATTTGGAAGAGTTTTCAGACCTGT 1920
Db 1861 GTGAAAGATTGATCAGTTCGCTGTTCGAAACATGTGATTTGGAAGAGTTTTCAGACCTGT 1920
Qy 1921 CTGCTGGGCACTGCAAGCATGGTTTGGCTGTGCTGTCGACACTCTGGCTGGAAA 1980
Db 1921 CTGCTGGGCACTGCAAGCATGGTTTGGCTGTGCTGTCGACACTCTGGCTGGAAA 1980
Qy 1981 GTGCTCTATTCCGGGGACACCATCCCTCGAGGCTCTGGTCCGATGGGAAAGATGCC 2040
Db 1981 GTGCTCTATTCCGGGGACACCATCCCTCGAGGCTCTGGTCCGATGGGAAAGATGCC 2040
Qy 2041 ACCCTCTGATACATGAAGCCACCTTGGAAAGATGGTTTGGAAAGAGGAGTCAAG 2100
Db 2041 ACCCTCTGATACATGAAGCCACCTTGGAAAGATGGTTTGGAAAGAGGAGTCAAG 2100
Qy 2101 ACACAGCACAAGCTCCCAAGCCATCAGCGTGGGATGGGATGAACGGGAGTTCA 2160
Db 2101 ACACAGCACAAGCTCCCAAGCCATCAGCGTGGGATGGGATGAACGGGAGTTCA 2160
Qy 2161 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGTCCCTCTTTCAGCCCCCACTCAGC 2220
Db 2161 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGTCCCTCTTTCAGCCCCCACTCAGC 2220
Qy 2221 GAGAAAGTGGAGTGTCTTTGACCATGAAGTCTGCTTTGGAGACTTTTCCAACTG 2280
Db 2221 GAGAAAGTGGAGTGTCTTTGACCATGAAGTCTGCTTTGGAGACTTTTCCAACTG 2280

QY 2281 CCCAAGCTGATTTCCCTCCCTGAAAGCCCTGTTGCTGGGACATCGAGAGATGGAGG 2340
DB 2281 CCCAAGCTGATTTCCCTCCCTGAAAGCCCTGTTGCTGGGACATCGAGAGATGGAGG 2340
QY 2341 CGCAGGAGAAAGCGGAGCTGCGGAGGTGCGGGCGGCCCTCTGTCACAGGAGCTGGCA 2400
DB 2341 CGCAGGAGAAAGCGGAGCTGCGGAGGTGCGGGCGGCCCTCTGTCACAGGAGCTGGCA 2400
QY 2401 GCGGGCTCGAGATGGGAGCTTACAGAGAGCGGGCCCAACAGAGAGCCACAGGCC 2460
DB 2401 GCGGGCTCGAGATGGGAGCTTACAGAGAGCGGGCCCAACAGAGAGCCACAGGCC 2460
QY 2461 AAGAAGTCCAGAGCCCAAGTGA 2481
DB 2461 AAGAAGTCCAGAGCCCAAGTGA 2481

RESULT 2

US-09-988-687-1
; Sequence 1, Application US/09988697
; Publication No. US20030045704A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,687
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2481
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
US-09-988-687-1

Query Match 100.0%; Score 2481; DB 9; Length 2481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTGGGCGCTTTGCTCGCTGCTGGGTCCGGCCGGGCGGACGACCATGTCGAGGAGCG 60
DB 1 ATGTGGGCGCTTTGCTCGCTGCTGGGTCCGGCCGGGCGGACGACCATGTCGAGGAGCG 60
QY 61 ACCATATCGAGGACCCCGCCCGCGAGCGCGCGCAAGACCGCTGCGGACCTG 120
DB 61 ACCATATCGAGGACCCCGCCCGCGAGCGCGCGCAAGACCCGCTGCGGACCTG 120
QY 121 CGCAGCGAGAGAAGCGCGGACCGTCCGGGTGCTCCGGCGGCCCAACACCGGTGTACCTG 180
DB 121 CGCAGCGAGAGAAGCGCGGACCGTCCGGGTGCTCCGGCGGCCCAACACCGGTGTACCTG 180
QY 181 CAGGTGGTGGACGGGTAGCGGAGTCCGGGCGCGCGCTCTAGCTTCTCCAGATTC 240
DB 181 CAGGTGGTGGACGGGTAGCGGAGTCCGGGCGCGCGCTCTAGCTTCTCCAGATTC 240
QY 241 AACCGGTATCTCTTCAACTGTGAGAGAGCGTTTCAGAGACTCATCAGAGCACAGTTA 300
DB 241 AACCGGTATCTCTTCAACTGTGAGAGAGCGTTTCAGAGACTCATCAGAGCACAGTTA 300

QY 301 AAGTTGCTCGCTCGGACAAACATATTTCTTGACACGAATGCACTGGTCTTAATGTTGGGGC 360
DB 301 AAGTTGCTCGCTCGGACAAACATATTTCTTGACACGAATGCACTGGTCTTAATGTTGGGGC 360
QY 361 TTAAGTGAATGATTTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTAATTTCTGGA 420
DB 361 TTAAGTGAATGATTTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTAATTTCTGGA 420
QY 421 CTTCCACAACCTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGA 480
DB 421 CTTCCACAACCTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGA 480
QY 481 ATAGAACTGGCTGTGCGGCCCTCTGCTCCAGAAATACGAGGATGAAACCAATGACAGTT 540
DB 481 ATAGAACTGGCTGTGCGGCCCTCTGCTCCAGAAATACGAGGATGAAACCAATGACAGTT 540
QY 541 TACCAGATCCCCATACACAGTGAACAGAGGAGGGAAGCAACCAATGCGAGAGTCCA 600
DB 541 TACCAGATCCCCATACACAGTGAACAGAGGAGGGAAGCAACCAATGCGAGAGTCCA 600
QY 601 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCCAGACTCCGAGTCCGAATGAAAT 660
DB 601 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCCAGACTCCGAGTCCGAATGAAAT 660
QY 661 GAGCCACACTTCCACATGTTAGCCAGAGAGAGGGGTGAGGACTCTTCCCTGGTC 720
DB 661 GAGCCACACTTCCACATGTTAGCCAGAGAGAGGGGTGAGGACTCTTCCCTGGTC 720
QY 721 GTAGCTTTCACTGTAAGCTTCACTTTAAGAGAGAACTCTTGGTCTCAAGCAAG 780
DB 721 GTAGCTTTCACTGTAAGCTTCACTTTAAGAGAGAACTCTTGGTCTCAAGCAAG 780
QY 781 GAGATGGGCTCCCAAGTGGGACAGCTGCAATCGCTCCCATCATTTGCTGCTCTCAAGGAC 840
DB 781 GAGATGGGCTCCCAAGTGGGACAGCTGCAATCGCTCCCATCATTTGCTGCTCTCAAGGAC 840
QY 841 GGGAAAGCATCACTCATGAAGGAGAGAGATTTGGCTGAAGAGCTGTGTACTCTCTCA 900
DB 841 GGGAAAGCATCACTCATGAAGGAGAGAGATTTGGCTGAAGAGCTGTGTACTCTCTCA 900
QY 901 GATCTGCTGCTGCTTTTGTGGTGTAGAAATGTCAGATGAAGCTTCAATCAACCCATC 960
DB 901 GATCTGCTGCTGCTTTTGTGGTGTAGAAATGTCAGATGAAGCTTCAATCAACCCATC 960
QY 961 TGTGAGAAATGCCACCTTTCCAGAGTACCAAGGAGGAGATGCCCGCTGGCTTGGTG 1020
DB 961 TGTGAGAAATGCCACCTTTCCAGAGTACCAAGGAGGAGATGCCCGCTGGCTTGGTG 1020
QY 1021 GTTCAATGGGCCCCAGCATCTGTGCTTGTGGACAGCAGGTACACAGTGGATGAGAGG 1080
DB 1021 GTTCAATGGGCCCCAGCATCTGTGCTTGTGGACAGCAGGTACACAGTGGATGAGAGG 1080
QY 1081 TTTGGGCTGACACCCAGCATTTGCTCTGATGAGAACTGTGCTCAGTTCAGTTCAACCTT 1140
DB 1081 TTTGGGCTGACACCCAGCATTTGCTCTGATGAGAACTGTGCTCAGTTCAGTTCAACCTT 1140
QY 1141 CGCAGCCACAAGATTCAAAACCCAGCTCAACCTCATCCACCGGACATCTTCCCCCTGCTC 1200
DB 1141 CGCAGCCACAAGATTCAAAACCCAGCTCAACCTCATCCACCGGACATCTTCCCCCTGCTC 1200
QY 1201 ACCAGTTTCCGCTGTAAGAGAGGGCCCCACCTCAGTGTGCCCATGTTTCAGGGTGA 1260
DB 1201 ACCAGTTTCCGCTGTAAGAGAGGGCCCCACCTCAGTGTGCCCATGTTTCAGGGTGA 1260
QY 1261 TGCCTCTCAAGTACAGCTCCGTCAGGAGGAGTGGCAGAGGATGCCATTATTACT 1320
DB 1261 TGCCTCTCAAGTACAGCTCCGTCAGGAGGAGTGGCAGAGGATGCCATTATTACT 1320
QY 1321 TGCATCTCTAGGAAATTCATAGTTGAGGCGGTGCAAGCTTCCCACTTCAGCAGAGCGTG 1380
DB 1321 TGCATCTCTAGGAAATTCATAGTTGAGGCGGTGCAAGCTTCCCACTTCAGCAGAGCGTG 1380
QY 1381 CAGGAGTACAGGAGGTGCGCAGGACGCGCCCGAGGAGAGAGAGAGAGAGTACAGTAC 1440

Db 1381 CAGGAGTACAGGAGGTGCGCAGACGGCCAGCCCGCAGACGAGAGAAAGAGTCAGTAC 1440
Qy 1441 CCAGAAATCATCTTCCCTTGGAAACAGGGCTGCGCATCCCGATGAAGATTGCGAAATGTCAGT 1500
Db 1441 CCAGAAATCATCTTCCCTTGGAAACAGGGCTGCGCATCCCGATGAAGATTGCGAAATGTCAGT 1500
Qy 1501 GCCACACTGTCAACATGAAGCCCGACAGCTCTGCTACTGGACTGTGGTGGGGCACA 1560
Db 1501 GCCACACTGTCAACATGAAGCCCGACAGCTCTGCTACTGGACTGTGGTGGGGCACA 1560
Qy 1561 TTTGGGCAGCTGTGCGGTCATTACGAGACACAGGTGGACAGGGTCTCGGCACCCCTGCT 1620
Db 1561 TTTGGGCAGCTGTGCGGTCATTACGAGACACAGGTGGACAGGGTCTCGGCACCCCTGCT 1620
Qy 1621 GCTGTGTTTGTGCCACTGACAGATCACACACGGGCTTCCCAAGTATCTTGTCTG 1680
Db 1621 GCTGTGTTTGTGCCACTGACAGATCACACACGGGCTTCCCAAGTATCTTGTCTG 1680
Qy 1681 CAGAGAGAACGGCCCTTGGCATCTTTTGGGAAAGCGCTTACCCCTTGTGTTGTTGCC 1740
Db 1681 CAGAGAGAACGGCCCTTGGCATCTTTTGGGAAAGCGCTTACCCCTTGTGTTGTTGCC 1740
Qy 1741 CCCAACACAGCTCAAGAGCTGGCTCCAGCAGTACCAACACAGTGCAGAGTCTCTGCAC 1800
Db 1741 CCCAACACAGCTCAAGAGCTGGCTCCAGCAGTACCAACACAGTGCAGAGTCTCTGCAC 1800
Qy 1801 CACATCAGTATGATTCCTGCCAAATGCTTCCAGGAGGGCTGAGATCTCCAGTCCCTGCA 1860
Db 1801 CACATCAGTATGATTCCTGCCAAATGCTTCCAGGAGGGCTGAGATCTCCAGTCCCTGCA 1860
Qy 1861 GTGGAAGATTGATCAGTTCGCTGTGGAACATGTGATTTGGAAGAGTTTCAGACCTGT 1920
Db 1861 GTGGAAGATTGATCAGTTCGCTGTGGAACATGTGATTTGGAAGAGTTTCAGACCTGT 1920
Qy 1921 CTGTTGGCGGCTCAAGCATGCTGCTTGGCTGTGCTGGTGCACACCTCTGGCTGGAAA 1980
Db 1921 CTGTTGGCGGCTCAAGCATGCTGTTGGCTGTGCTGGTGCACACCTCTGGCTGGAAA 1980
Qy 1981 GTGTTCTATTCCGGGACACCATGCTTGGTGGGATGCGGATGGGAAAGATGCC 2040
Db 1981 GTGTTCTATTCCGGGACACCATGCTTGGTGGGATGCGGATGGGAAAGATGCC 2040
Qy 2041 ACCCTCCTGATACATGAAGCCACCTTGGAGATGTTTGGAGAGAGCAGTGGAAAG 2100
Db 2041 ACCCTCCTGATACATGAAGCCACCTTGGAGATGTTTGGAGAGAGCAGTGGAAAG 2100
Qy 2101 ACACACAGCACAACTGCCAAGCCATCAGCGTGGGATGCGGATGAACCGCGAGTTCAAT 2160
Db 2101 ACACACAGCACAACTGCCAAGCCATCAGCGTGGGATGCGGATGAACCGCGAGTTCAAT 2160
Qy 2161 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGTCCCGCTTCCAGCCCAACTTCAGC 2220
Db 2161 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGTCCCGCTTCCAGCCCAACTTCAGC 2220
Qy 2221 GAGAAAGTGGAGTTGCTTTGACACATGAAGTCTGCTTTGGAGACTTTTCCAAATG 2280
Db 2221 GAGAAAGTGGAGTTGCTTTGACACATGAAGTCTGCTTTGGAGACTTTTCCAAATG 2280
Qy 2281 CCCAAGCTGATTTCCCCCACTGAAGCCCTGTTTGTGGCGACATCAGAGATGGAGAG 2340
Db 2281 CCCAAGCTGATTTCCCCCACTGAAGCCCTGTTTGTGGCGACATCAGAGATGGAGAG 2340
Qy 2341 CGCAGGAGAAAGCGGAGCTGCGGAGGTGCGGGCGGCCCTTCTTCCAGGAGCTGGCA 2400
Db 2341 CGCAGGAGAAAGCGGAGCTGCGGAGGTGCGGGCGGCCCTTCTTCCAGGAGCTGGCA 2400
Qy 2401 GCGGCCCTGGAGGATGGGAGCTCAGCAGAGAGCGGGGCCACACAGAGAGCCACAGGCC 2460
Db 2401 GCGGCCCTGGAGGATGGGAGCTCAGCAGAGAGCGGGGCCACACAGAGAGCCACAGGCC 2460
Qy 2461 AAGAAAGTCAGAGGCCAGTGA 2481

Db 2461 AAGAAAGTCAGAGGCCAGTGA 2481
RESULT 3
US-09-988-626-3
; Sequence 3, Application US/09988626
; Publication No. US20030044959A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,626
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2958
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51)..(2531)
; OTHER INFORMATION: coding sequence as in SEQ ID NO:1
US-09-988-626-3

Query Match 100.0%; Score 2481; DB 9; Length 2958;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGTGGGCGCTTTGCTCGCTGCTCGGTCGCGGCGCGGACGACCATGTCTCGAGGAGCG 60
Db 51 ATGTGGGCGCTTTGCTCGCTGCTCGGTCGCGGCGCGGACGACCATGTCTCGAGGAGCG 110
Qy 61 ACCATATCGAGGACCGCGCGCGCGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 120
Db 111 ACCATATCGAGGACCGCGCGCGCGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 170
Qy 121 CGCACGCGAGAGCGCGGACCGTCCGGGGTCTCCGGCGGCCCAACACCGCTGTACCTG 180
Db 171 CGCACGCGAGAGCGCGGACCGTCCGGGGTCTCCGGCGGCCCAACACCGCTGTACCTG 230
Qy 181 CAGGTGTGGCAGCGGGTAGCGGGGACTCGGGCGCGCGGCTCTACGCTCTTCTCCGAGTTC 240
Db 231 CAGGTGTGGCAGCGGGTAGCGGGGACTCGGGCGCGCGGCTCTACGCTCTTCTCCGAGTTC 290
Qy 241 AACCGGTATCTCTTCAACTGTGGAGAGGGCTTCAGAGACTCATGCGAGGAGCACAAGTTA 300
Db 291 AACCGGTATCTCTTCAACTGTGGAGAGGGCTTCAGAGACTCATGCGAGGAGCACAAGTTA 350
Qy 301 AAGGTTCTCGCTGGGACACATATTCCTGACACGAATGCACTGCTTAATGTTGGGGGC 360
Db 351 AAGGTTCTCGCTGGGACACATATTCCTGACACGAATGCACTGCTTAATGTTGGGGGC 410
Qy 361 TTAAGTGAATGATCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGA 420
Db 411 TTAAGTGAATGATCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGA 470
Qy 421 CCTCCACAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGA 480
Db 471 CCTCCACAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGA 530
Qy 481 ATGAACTGGGTGTGCGGCCCCACTCTGCCCCAGAAATACGAGGATGAACCATGACACTT 540

Db 531 ATAGAACTGGCTGTGGGCCCCCACTCTGCCCCAGAAATACGAGGATGAACCAATGACAGTT 590
Qy 541 TACCAGATCCCATACACAGTGAACAGAGAGGGGAAAGCACCACCAATGGCAGAGTCCA 600
Db 591 TACCAGATCCCATACACAGTGAACAGAGAGGGGAAAGCACCACCAATGGCAGAGTCCA 650
Qy 601 GAAAGGCTCTCAGCGAGGTCTAGTCAGAGCGATCTTACAGTCCGAGTCGAATAAT 660
Db 651 GAAAGGCTCTCAGCAGGCTCAGTCAGAGCGATCTTACAGTCCGAGTCGAATAAT 710
Qy 661 GAGCACACCTTCCACATGGTGTAGCCAGAGAGAGGGGTGAGGACCTTCCCTGGTC 720
Db 711 GAGCACACCTTCCACATGGTGTAGCCAGAGAGAGGGGTGAGGACCTTCCCTGGTC 770
Qy 721 GTAGCTTTTCATCTGAAGCTTTCACCTTAAAGAGAGAAACTTCTGGTGTCTCAAGCAAG 780
Db 771 GTAGCTTTTCATCTGAAGCTTTCACCTTAAAGAGAGAAACTTCTGGTGTCTCAAGCAAG 830
Qy 781 GAGATGGGCTCCAGTTGGGACAGTCGCCATCGCTCCCATATTGCTGCTGTCAAGGAC 840
Db 831 GAGATGGGCTCCAGTTGGGACAGTCGCCATCGCTCCCATATTGCTGCTGTCAAGGAC 890
Qy 841 GGGAAAGCATCACTCATGAAGGAGAGAGATTTGGCTGAAGAGCTGTGTACTCTCTCCA 900
Db 891 GGGAAAGCATCACTCATGAAGGAGAGAGATTTGGCTGAAGAGCTGTGTACTCTCTCCA 950
Qy 901 GATCTGGTGTCTTTTGTGGTGTAGATGTCCAGATGAAAGCTTCAATCAACCCCATC 960
Db 951 GATCTGGTGTCTTTTGTGGTGTAGATGTCCAGATGAAAGCTTCAATCAACCCCATC 1010
Qy 961 TGTGAGATGCGACCTTTTCAAGGATCAAGAAAGGAGATGCCCTGGCTGGCTGGTG 1020
Db 1011 TGTGAGATGCGACCTTTTCAAGGATCAAGAAAGGAGATGCCCTGGCTGGCTGGTG 1070
Qy 1021 GTTCATGGGCCCCAGCATCTGTGCTTGTGGACAGAGGTACAGAGTGGATGGAGG 1080
Db 1071 GTTCATGGGCCCCAGCATCTGTGCTTGTGGACAGAGGTACAGAGTGGATGGAGG 1130
Qy 1081 TTTGGGCTTGACACCCAGCAGTGTGCTTGTGAATGAAATGTGCTCAGTTCACAACTT 1140
Db 1131 TTTGGGCTTGACACCCAGCAGTGTGCTTGTGAATGAAATGTGCTCAGTTCACAACTT 1190
Qy 1141 CGAGGCCAAGATTAACCCAGTCAACCTCATCCACCGGACATCTTCCCTGGTC 1200
Db 1191 CGAGGCCAAGATTAACCCAGTCAACCTCATCCACCGGACATCTTCCCTGGTC 1250
Qy 1201 ACCAGTTTCCGCTGTGAAGAGAGGGCCCCCAGTGTGCCATGTTTCAAGGTGA 1260
Db 1251 ACCAGTTTCCGCTGTGAAGAGAGGGCCCCCAGTGTGCCATGTTTCAAGGTGA 1310
Qy 1261 TGCTCTCTCAAGTACAGCTCCGTCAGAGAGGAGTGGCAGAGGATGCCATTATTACT 1320
Db 1311 TGCTCTCTCAAGTACAGCTCCGTCAGAGAGGAGTGGCAGAGGATGCCATTATTACT 1370
Qy 1321 TGCAATCTTGAGGAATTCATAGTTGAGGCGCTGACGCTTCCCACTTCCAGCAGAGCGTG 1380
Db 1371 TGCAATCTTGAGGAATTCATAGTTGAGGCGCTGACGCTTCCCACTTCCAGCAGAGCGTG 1430
Qy 1381 CAGGAGTACAGAGAGTGGCAGAGCGGCCCGCCAGCCCGCCAGAGAAAGTCAAGTAC 1440
Db 1431 CAGGAGTACAGAGAGTGGCAGAGCGGCCCGCCAGCCCGCCAGAGAAAGTCAAGTAC 1490
Qy 1441 CCAGAAATCATCTTCTTGGAAAGAGGCTGTGCCATCCCATGAAGATTCGAAATGTCACT 1500
Db 1491 CCAGAAATCATCTTCTTGGAAAGAGGCTGTGCCATCCCATGAAGATTCGAAATGTCACT 1550
Qy 1501 GCCACATTTGTCAACATAAGCCCCGACAGCTCTCTGCTACTGAGCTGTGGTGGAGGCA 1560
Db 1551 GCCACATTTGTCAACATAAGCCCCGACAGCTCTCTGCTACTGAGCTGTGGTGGAGGCA 1610
Qy 1561 TTTGGCAGCTGTGCGCTCATTAACGAGACCAAGGTGGAAGGGTCTGGGACCTTGGCT 1620

Db 1611 TTTGGCAGCTGTGCGCTCATTAACGAGACCAAGGTGGACAGGGTCTTGGGACACCCTGGCT 1670
Qy 1621 GCTGTGTTTGTGTCCACCTGCACGACAGATCACACACGGGCTTCCCAAGTATCTTGTG 1680
Db 1671 GCTGTGTTTGTGTCCACCTGCACGACAGATCACACACGGGCTTCCCAAGTATCTTGTG 1730
Qy 1681 CAGAGAGAACGGCTTGGCATCTTTGGGAAAGCGCTTCAACCTTCTGCTGGTGGTGGCC 1740
Db 1731 CAGAGAGAACGGCTTGGCATCTTTGGGAAAGCGCTTCAACCTTCTGCTGGTGGTGGCC 1790
Qy 1741 CCCAACACAGCTCAAAAGCTGCTCCAGCAGTACCAACACAGTGCAGAGGTCTCTGCAC 1800
Db 1791 CCCAACACAGCTCAAAAGCTGCTCCAGCAGTACCAACACAGTGCAGAGGTCTCTGCAC 1850
Qy 1801 CACATCAGTATGATTCCTGCGCAAAATGCCCTTCAAGAAAGGGCTGAGATCTCCAGTCTTCCA 1860
Db 1851 CACATCAGTATGATTCCTGCGCAAAATGCCCTTCAAGAAAGGGCTGAGATCTCCAGTCTTCCA 1910
Qy 1861 GTGNAAGATTGATCAGTTCGCTGTGCGAATGTGATTTGGAAGAGTTTCAGACCTGT 1920
Db 1911 GTGNAAGATTGATCAGTTCGCTGTGCGAATGTGATTTGGAAGAGTTTCAGACCTGT 1970
Qy 1921 CTGTGCGGCACTGCAAGCATGCTTTTGGCTGTGCGCTGGTGCACACCTCTGGCTGGAAA 1980
Db 1971 CTGTGCGGCACTGCAAGCATGCTTTTGGCTGTGCGCTGGTGCACACCTCTGGCTGGAAA 2030
Qy 1981 GTGTCTTATTCCGGGACACCATGCCCTCGCAGGCTCTGGTCCGGATGGGAAAGATGCC 2040
Db 2031 GTGTCTTATTCCGGGACACCATGCCCTCGCAGGCTCTGGTCCGGATGGGAAAGATGCC 2090
Qy 2041 ACCCTCTGATACATGAAGCCACCTCGAAGATGTTTGGAAAGAGGACAGTGGAAAAG 2100
Db 2091 ACCCTCTGATACATGAAGCCACCTCGAAGATGTTTGGAAAGAGGACAGTGGAAAAG 2150
Qy 2101 ACACACACACAACTCCCAAGCCATCAGCGTGGGAGTCCGATGAAACGCGAGTTTCA 2160
Db 2151 ACACACACACAACTCCCAAGCCATCAGCGTGGGAGTCCGATGAAACGCGAGTTTCA 2210
Qy 2161 ATGCTGAACCACTTTCAGCCAGCGCTATCCAAAGTTCCTTTCAGCCCCCACTTCA 2220
Db 2211 ATGCTGAACCACTTTCAGCCAGCGCTATCCAAAGTTCCTTTCAGCCCCCACTTCA 2270
Qy 2221 GAGAAAGTGGAGTTGCTTTGACACATGAAGTCTGCTTTGGAGACTTTTCCAAATG 2280
Db 2271 GAGAAAGTGGAGTTGCTTTGACACATGAAGTCTGCTTTGGAGACTTTTCCAAATG 2330
Qy 2281 CCCAAGCTGATTTCCCACTGAAAGCCCTGTTGCTGCGACATCGAGAGATGGAGG 2340
Db 2331 CCCAAGCTGATTTCCCACTGAAAGCCCTGTTGCTGCGACATCGAGAGATGGAGG 2390
Qy 2341 CGCAGGAGAAAGCGGAGCTGCGGAGGTGGGCGGCGCTTCTTCCAGGAGCTGGCA 2400
Db 2391 CGCAGGAGAAAGCGGAGCTGCGGAGGTGGGCGGCGCTTCTTCCAGGAGCTGGCA 2450
Qy 2401 GCGGCTTGGAGATGGGAGCTTCAAGAGCGGCTTCAAGAGCGGCGCCACACAGAGAGCCA 2460
Db 2451 GCGGCTTGGAGATGGGAGCTTCAAGAGCGGCTTCAAGAGCGGCGCCACACAGAGAGCCA 2510
Qy 2461 AAGAGGTTCAGAGCCCACTGA 2481
Db 2511 AAGAGGTTCAGAGCCCACTGA 2531

RESULT 4

US-09-988-687-3
; Sequence 3, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigan, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.


```
Qy 901 GATCCTGCTGCTGCTTTTGTGTGTAGTAATGTCAGATGAAAGCTTCATTCAACCCATC 960
Db 901 GATCCTGCTGCTGCTTTTGTGTGTGTAGTAATGTCAGATGAAAGCTTCATTCAACCCATC 960
Qy 961 TGTGGAATGCAACCTTTTCAGAGTACCAAGAAAGGAGAGATGCCCCGTCCTGTTGGT 1020
Db 961 TGTGGAATGCAACCTTTTCAGAGTACCAAGAAAGGAGAGATGCCCCGTCCTGTTGGT 1020
Qy 1021 GTTCATATGGCCCCAGCATCTGTGCTTGTGGACACAGGTACACAGTGGATGAGAGG 1080
Db 1021 GTTCATATGGCCCCAGCATCTGTGCTTGTGGACACAGGTACACAGTGGATGAGAGG 1080
Qy 1081 TTTGGGCTGACACCCAGCACTTGTGCTGTGAATGAGAACTGTGCTCAGTTTCAACAACCTT 1140
Db 1081 TTTGGGCTGACACCCAGCACTTGTGCTGTGAATGAGAACTGTGCTCAGTTTCAACAACCTT 1140
Qy 1141 CGACGCCAAGATTTCAACCCAGCTCAACCTTCATCCACCGGACATCTTCCCCCTGCTC 1200
Db 1141 CGACGCCAAGATTTCAACCCAGCTCAACCTTCATCCACCGGACATCTTCCCCCTGCTC 1200
Qy 1201 ACCAGTTTCCGCTTAAGAAAGAGGGCCCCACCCCTCAGTGTGCCATGGTTCAAGGTGAA 1260
Db 1201 ACCAGTTTCCGCTTAAGAAAGAGGGCCCCACCCCTCAGTGTGCCATGGTTCAAGGTGAA 1260
Qy 1261 TGCTCTCTCAAGTACAGCTCCGCTCCAGAGGGAGTGGCAGAGGATGCCATTTACT 1320
Db 1261 TGCTCTCTCAAGTACAGCTCCGCTCCAGAGGGAGTGGCAGAGGATGCCATTTACT 1320
Qy 1321 TGCATCTCTGAGGAATTCATAGTTGAGGCGTTCAGCTTCCAACTTCACGAGAGCGTG 1380
Db 1321 TGCATCTCTGAGGAATTCATAGTTGAGGCGTTCAGCTTCCAACTTCACGAGAGCGTG 1380
Qy 1381 CAGGAGTACAGAGAGTGGCAGAGCGGCCAGCCCGAGCCAGCAGAGAAAGAGTCAGTAC 1440
Db 1381 CAGGAGTACAGAGAGTGGCAGAGCGGCCAGCCCGAGCCAGCAGAGAAAGAGTCAGTAC 1440
Qy 1441 CCAGAAATCATCTTCTTGGAAACAGGCTCTGCCATCCCGATCCCGATTCGAAATGTCAGT 1500
Db 1441 CCAGAAATCATCTTCTTGGAAACAGGCTCTGCCATCCCGATTCGAAATGTCAGT 1500
Qy 1501 GCCACACTTGTCAACATAAGCCCCGACACGCTCTCTGCTACTGGAATGTTGGTGGAGGACA 1560
Db 1501 GCCACACTTGTCAACATAAGCCCCGACACGCTCTCTGCTACTGGAATGTTGGTGGAGGACA 1560
Qy 1561 TTTGGGCGAGCTGTCGCTCATTTACGAGACACAGTGGAGAGGCTCTGGGCACCCCTGCT 1620
Db 1561 TTTGGGCGAGCTGTCGCTCATTTACGAGACACAGTGGAGAGGCTCTGGGCACCCCTGCT 1620
Qy 1621 GCTGTGTTGTGTCCTCCACCTGCACGAGATCACACACGGGCTTGCAGATCTTGTGCTG 1680
Db 1621 GCTGTGTTGTGTCCTCCACCTGCACGAGATCACACACGGGCTTGCAGATCTTGTGCTG 1680
Qy 1681 CAGAGAGAACCGCCTTGGCATCTTTTGGAAAGCCGCTTCAACCTTGTGCTGTTGCC 1740
Db 1681 CAGAGAGAACCGCCTTGGCATCTTTTGGAAAGCCGCTTCAACCTTGTGCTGTTGCC 1740
Qy 1741 CCCAACCCAGCTCAAGCCCTGGCTCCAGCAGTACCAACACAGTGGCAGAGGTCCTGCAC 1800
Db 1741 CCCAACCCAGCTCAAGCCCTGGCTCCAGCAGTACCAACACAGTGGCAGAGGTCCTGCAC 1800
Qy 1801 CACATCAGTATGATTTCTCCCAAATGCCCTTCAGGAAGGGGCTGAGATCTCCAGTCTCTCA 1860
Db 1801 CACATCAGTATGATTTCTCCCAAATGCCCTTCAGGAAGGGGCTGAGATCTCCAGTCTCTCA 1860
Qy 1861 GTGGAAGATTCATCAGTTTCGCTGTTGGCAACATGTGATTTTGGAAAGTTTCAGACCTGT 1920
Db 1861 GTGGAAGATTCATCAGTTTCGCTGTTGGCAACATGTGATTTTGGAAAGTTTCAGACCTGT 1920
Qy 1921 CTGGTGGCGCACTGAAGCATCGTTTGGCTGTGGCTGGTGCACACCTCTCGCTGGGAAA 1980
Db 1921 CTGGTGGCGCACTGAAGCATCGTTTGGCTGTGGCTGGTGCACACCTCTCGCTGGGAAA 1980
Qy 1981 GTGGTCTATTCCGGGACACCATGCCCCTCGAGGCTCTGGTCCGGATGGGAAAGATGCC 2040
```

```
Db 1981 GTGCTCTATTCCGGGGACACCATGCCCTCGAGGCTCTGTGTCGGATGGGAAAAGATGCC 2040
Qy 2041 ACCCTCTCTGATACATGAAGCCACCCTCGAAGATGTTTGGNAGAGGAGCAGTGGAAAAG 2100
Db 2041 ACCCTCTCTGATACATGAAGCCACCCTCGAAGACGTTTGGNAGAGGAGCAGTGGAAAAG 2100
Qy 2101 ACACACAGCAACGCTCCCAAGCCATCAGCGTGGGATGCGGATGAACGCGGAGTTCAATT 2160
Db 2101 ACACACAGCAACGCTCCCAAGCCATCAGCGTGGGATGCGGATGAACGCGGAGTTCAATT 2160
Qy 2161 ATGCTGAACCACTTTCAGCCAGCGCTATGCCAAAGTCCCTCTTTTCAGCCCCCACTTCAAC 2220
Db 2161 ATGCTGAACCACTTTCAGCCAGCGCTATGCCAAAGTCCCTCTTTTCAGCCCCCACTTCAAC 2220
Qy 2221 GAGAAAGTGGGAGTTGCTTTGACCACATGAAGTCTGCTTTGAGACTTTTGCACAAATG 2280
Db 2221 GAGAAAGTGGGAGTTGCTTTGACCACATGAAGTCTGCTTTGAGACTTTTGCACAAATG 2280
Qy 2281 CCCAAGCTGATTTCCCCACTGAAAGCCCTGTTTCTGGCAGATCGAGGAGATGGAGAG 2340
Db 2281 CCCAAGCTGATTTCCCCACTGAAAGCCCTGTTTCTGGCAGATCGAGGAGATGGAGAG 2340
Qy 2341 CGCAGGAGAGAGCGGAGCTTCGCGCAGTTCGCGCGGCGCCCTCTCTTCCAGGAGCTGGCA 2400
Db 2341 CGCAGGAGAGAGCGGAGCTTCGCGCAGTTCGCGCGGCGCCCTCTCTTCCAGGAGCTGGCA 2400
Qy 2401 GCGCGCTGAGGATGGGAGCCTCAGCAGAAAGGGGCCCCACACAGAGGAGCCACAGGCC 2460
Db 2401 GCGCGCTGAGGATGGGAGCCTCAGCAGAAAGGGGCCCCACACAGAGGAGCCACAGGCC 2460
Qy 2461 AAGAAAGTTCAGAGCCCACTGA 2481
Db 2461 AAGAAAGTTCAGAGCCCACTGA 2481
```

RESULT 6

```
US-09-988-687-223
; Sequence 223, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
; APPLICANT: Tavtighian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,687
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/584,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 223
; LENGTH: 2908
; TYPE: DNA
; ORGANISM: Pan troglodytes
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
US-09-988-687-223
```

```
Query Match 99.0%; Score 2455.4; DB 9; Length 2908;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 2465; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
Qy 1 ATGTGGCGCTTTGCTCGCTGCTGCGGTCCGCGGCGGACGACCACTGTCGAGGGAGCGC 60
```

Db 1 ATGTGGCGCTTTGCTCGTGTCTGGTCCGGCCGGACGACCATGTCCAGGAGCC 60
Qy 61 ACCATATCCAGGCAACCCCGCCCGCAGCGCGCGAAGGACCCGCTGGGCACTG 120
Db 61 ACCATATCCAGGCAACCCCGCCCGCAGCGCGCGAAGGACCCGCTGGGCACTG 120
Qy 121 CCACGCGAGAGAGCGGCGGACCGTGGGGTCTCGGGCGGCCAAACACCGTGTACCTG 180
Db 121 CCACGCGAGAGAGCGGCGGACCGTGGGGTCTCGGGCGGCCAAACACCGTGTACCTG 180
Qy 181 CAGGTGGTGGCAGCGGATAGCGGACCTCGGCGCGCGCTCTAGCTCTTCCGAGTTC 240
Db 181 CAGGTGGTGGCAGCGGATAGCGGACCTCGGCGCGCGCTCTAGCTCTTCCGAGTTC 240
Qy 241 AACCGGTATCTCTTCAACTGTGGAGAGCGCTTACAGAGACTCATGAGGAGCACAAGTTA 300
Db 241 AACCGGTATCTCTTCAACTGTGGAGAGCGCTTACAGAGACTCATGAGGAGCACAAGTTA 300
Qy 301 AAGGTGTCTCGCTGGACAAATATTCCTGACACGAAATGCACTGGTCTAATGTGGGGC 360
Db 301 AAGGTGTCTCGCTGGACAAATATTCCTGACACGAAATGCACTGGTCTAATGTGGGGC 360
Qy 361 TTAAGTGGAAATGATCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTACTTTCTGGA 420
Db 361 TTAAGTGGAAATGATCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTACTTTCTGGA 420
Qy 421 CTTCCACAACCTGGAAAAATPACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGA 480
Db 421 CTTCCACAACCTGGAAAAATPACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGA 480
Qy 481 ATAGAACTGGTGTGGGCGCCCACTCTGCCCAGAAATACAGAGATGAACACCATGACGTT 540
Db 481 ATAGAACTGGTGTGGGCGCCCACTCTGCCCAGAAATACAGAGATGAACACCATGACGTT 540
Qy 541 TACCGATCCCATACACAGTGAACAGAGAGGGGAAAGCAACCACTGGCAGAGTCCA 600
Db 541 TACCGATCCCATACACAGTGAACAGAGAGGGGAAAGCAACCACTGGCAGAGTCCA 600
Qy 601 GAAAGGCTCTCAGCAGGCTCAGTCAGAGCGATCTTTCAGACTCCGAGTCGAATGAAAT 660
Db 601 GAAAGGCTCTCAGCAGGCTCAGTCAGAGCGATCTTTCAGACTCCGAGTCGAATGAAAT 660
Qy 661 GAGCACACCTTCCATCGTGTAGCCAGAGAGAGGGGTGAGGACTCTTCCCTGGTC 720
Db 661 GAGCACACCTTCCATCGTGTAGCCAGAGAGAGGGGTGAGGACTCTTCCCTGGTC 720
Qy 721 GTAGCTTTTCATCTGAAGCTTCACTTAAGAGAGGAACTTCTGGTGTCTCAAGCAAG 780
Db 721 GTAGCTTTTCATCTGAAGCTTCACTTAAGAGAGGAACTTCTGGTGTCTCAAGCAAG 780
Qy 781 GAGATGGGCTCCAGTGGGACAGTCGCTCCATCTGCTCATCTGCTCAAGGAC 840
Db 781 GAGATGGGCTCCAGTGGGACAGTCGCTCCATCTGCTCATCTGCTCAAGGAC 840
Qy 841 GGGAAAGCATCACTCATGAAGAGAGAGATTTGGCTGAAGAGCTGTACTCTCTCA 900
Db 841 GGGAAAGCATCACTCATGAAGAGAGAGATTTGGCTGAAGAGCTGTACTCTCTCA 900
Qy 901 GATCTGTGTCTGCTTTGTGTGTGTAGATGTCCAGATGAAGCTTCAATCAACCCATC 960
Db 901 GATCTGTGTCTGCTTTGTGTGTGTAGATGTCCAGATGAAGCTTCAATCAACCCATC 960
Qy 961 TGTGAGATGCCACCTTTTCAGAGGTACCAAGGAAAGGAGATGCCCGTGGCTTGGTG 1020
Db 961 TGTGAGATGCCACCTTTTCAGAGGTACCAAGGAAAGGAGATGCCCGTGGCTTGGTG 1020
Qy 1021 GTTACATGGCCCCCAGCATCTGTGCTTGTGGAGCAGCAGGTACAGCAGTGGATGGAGAG 1080
Db 1021 GTTACATGGCCCCCAGAACTGTGCTTGTGGAGCAGCAGGTACAGCAGTGGATGGAGAG 1080
Qy 1081 TTTGGGCTTGACCCAGCACTTGTGCTCTGAATGAGAACTGTGCTCAGTTCAACACCTT 1140
Db 1081 TTTGGGCTTGACCCAGCACTTGTGCTCTGAATGAGAACTGTGCTCAGTTCAACACCTT 1140

Db 1081 TTTGGGCTTGACCCAGCACTTGTGCTCTGAATGAGAACTGTGCTCAGTTCAACACCTT 1140
Qy 1141 CGAGCCCAAGATTTCAAACCCAGCTCAACCTCATCCACCGGACATCTTCCCCCTGTC 1200
Db 1141 CGAGCCCAAGATTTCAAACCCAGCTCAACCTCATCCACCGGACATCTTCCCCCTGTC 1200
Qy 1201 ACCAGTTTCCGCTGTAAAGAGAGGGGCCCAACCTCAGTGTGCCCATGGTTCAGGTTGAA 1260
Db 1201 ACCAGTTTCCGCTGTAAAGAGAGGGGCCCAACCTCAGTGTGCCCATGGTTCAGGTTGAA 1260
Qy 1261 TGCCTCTCAAGTACCAGTCCGTCAGGAGGAGTGGCAGAGGATGCCATTATTACT 1320
Db 1261 TGCCTCTCAAGTACCAGTCCGTCAGGAGGAGTGGCAGAGGATGCCATTATTACT 1320
Qy 1321 TGCATCTCTGAGGAATTCATAGTTGAGCGCTGTCAGCTTCCCAACTTCAGCAGAGCGTG 1380
Db 1321 TGCATCTCTGAGGAATTCATAGTTGAGCGCTGTCAGCTTCCCAACTTCAGCAGAGTG 1380
Qy 1381 CAGGAGTACAGGAGTGGCAGGACCGCCAGCCCGAGCCCGAGCAGAGAGAAAGTCAGTAC 1440
Db 1381 CAGGAGTACAGGAGTGGCAGGACCGCCAGCAGCGCCCGAGCAGAGAGAAAGTCAGTAC 1440
Qy 1441 CCAGAAATCATCTTCTTTGGAACAGGGTCTGCATCCCGATGAAGATTGCAAAATGTCACT 1500
Db 1441 CCAGAAATCATCTTCTTTGGAACAGGGTCTGCATCCCGATGAAGATTGCAAAATGTCACT 1500
Qy 1501 GCCACATCTGTCAACATAAGCCCGACACGCTCTCTGTACTGGACTGTGGTGGAGGCACA 1560
Db 1501 GCCACATCTGTCAACATAAGCCCGACACGCTCTCTGTACTGGACTGTGGTGGAGGCAG 1560
Qy 1561 TTTGGGCACTGTGCGCTCATTAACGAGACAGGTGGAGAGGTCTCGGCAACCTGGCT 1620
Db 1561 TTTGGGCACTGTGCGCTCATTAACGAGACAGGTGGAGAGGTCTCGGCAACCTGGCT 1620
Qy 1621 GCTGTGTTGTGTCCTCCACTGCACGACATCACACAGGGCTTCCCAAGTATCTTGTG 1680
Db 1621 GCTGTGTTGTGTCCTCCACTGCACGACATCACACAGGGCTTCCCAAGTATCTTGTG 1680
Qy 1681 CAGAGAGAACCGCTTGGCATCTTTGGGAAAGCGCTTTCACCTTTGCTGGTGTGCC 1740
Db 1681 CAGAGAGAACCGCTTGGCATCTTTGGGAAAGCGCTTTCACCTTTGCTGGTGTGCC 1740
Qy 1741 CCCAACAGCTCAAAGCCTCGGCTCCAGCAGTACACACAGTGCAGAGGTCTCTGCAC 1800
Db 1741 CCCAACAGCTCAAAGCCTCGGCTCCAGCAGTACACACAGTGCAGAGGTCTCTGCAC 1800
Qy 1801 CACATCAGTATGATTCCTGCAAAATGCCTTCAGGAAGGGCTGAGATCTCCAGTCTTCA 1860
Db 1801 CACATCAGTATGATTCCTGCAAAATGCCTTCAGGAAGGGCTGAGATCTCCAGTCTTCA 1860
Qy 1861 GTGAAAGATTTGATCAGTTCGCTGTTCGAAACATGTGATTTGGAAGAGTTTCAGACCTGT 1920
Db 1861 GTGAAAGATTTGATCAGTTCGCTGTTCGAAACATGTGATTTGGAAGAGTTTCAGACCTGT 1920
Qy 1921 CTGTTGCGGCACTGCAAGCATGCGTTTGGCTGTGCGCTGGTGCACACCTCTGGCTGGAAA 1980
Db 1921 CTGTTGCGGCACTGCAAGCATGCGTTTGGCTGTGCGCTGGTGCACACCTCTGGCTGGAAA 1980
Qy 1981 GTGTTCTATTTCGGGGACACCATGCCCTCGAGGCTCTGGTCCGATGGGGAAGATGCC 2040
Db 1981 GTGTTCTATTTCGGGGACACCATGCCCTCGAGGCTCTGGTCCGATGGGGAAGATGCC 2040
Qy 2041 ACCCTCTGATACATGAAGCCACCTTGAAGCGGTTTGGAAAGAGAGTGGAAAG 2100
Db 2041 ACCCTCTGATACATGAAGCCACCTTGAAGCGGTTTGGAAAGAGAGTGGAAAG 2100
Qy 2101 ACACAGCACAACTCCCAAGCCATCAGCGTGGGATGCGGATGAACCGGAGTTCAAT 2160
Db 2101 ACACAGCACAACTCCCAAGCCATCAGCGTGGGATGCGGATGAACCGGAGTTCAAT 2160
Qy 2161 ATGCTGAACCACTTCAGCGAGGCTATGCCAAGTCCCGCTCTTCAGCCCCCACTTCAGC 2220
Db 2161 ATGCTGAACCACTTCAGCGAGGCTATGCCAAGTCCCGCTCTTCAGCCCCCACTTCAGC 2220

```
QY 2221 GAGAAAGTGGAGTTGCCTTTGACACATGAGGTCTGCTTTGGAGACTTTCCAAACATG 2280
Db 2221 GAGAAAGTGGAGTTGCCTTTGACACATGAGGTCTGCTTTGGAGACTTTGCAACAATG 2280
QY 2281 CCCAAGCTGATTCCCCCACTGAAAGCCCTGTTTGTGCGCATCGAGAGATGGAGAG 2340
Db 2281 CCCAAGCTGATTCCCCCACTGAAAGCCCTGTTTGTGCGCATCGAGAGATGGAGAG 2340
QY 2341 CGCAGGGAAGCGGAGTGGCGAGGTGGCGGCGGCCCTCTCTCCAGGAGCTGGCA 2400
Db 2341 CGCAGGGAAGCGGAGTGGCGAGGTGGCGGCGGCCCTCTCTCCAGGAGCTGGCA 2400
QY 2401 GCGCGCCTGGAGTGGGAGCCTCAGCAGAGCGGCGGCCACACAGAGAGCCACAGGCC 2460
Db 2401 GCGCGCCTGGAGTGGGAGCCTCAGCAGAGCGGCGGCCACACAGAGAGCCACAGGCC 2460
QY 2461 AAGAAAGTCAGAGCCAGTGA 2481
Db 2461 AAGAAAGTCAGAGCCAGTGA 2481

RESULT 7
US-09-988-626-225
; Sequence 225, Application US/09988626
; Publication No. US20030044959A1
; GENERAL INFORMATION:
; APPLICANT: Tavtighian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,626
; PRIORITY FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 225
; LENGTH: 2892
; TYPE: DNA
; ORGANISM: Gorilla gorilla
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
US-09-988-626-225

Query Match 98.5%; Score 2442.6; DB 9; Length 2892;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 2457; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 1 ATGTGGCGCTTTGCTCGCTGTCGGTCCGCGCGGAGCGCACCATTGTCGAGGAGCGC 60
Db 1 ATGTGGCGCTTTGCTCGCTGTCGGTCCGCGCGGAGCGCACCATTGTCGAGGAGCGC 60
QY 61 ACCATATCGCAGGACCCGCGCGGAGCGCGCGCGCGCAGGACCCGCTCGGACCTG 120
Db 61 ACCATATCGCAGGACCCGCGCGGAGCGCGCGCGCGCAGGACCCGCTCGGACCTG 120
QY 121 CGCAGCGAGAGAGCGCGGACCGTCGGGTGCTCCGCGCGCCAAACACCGTGTACCTG 180
Db 121 CGCAGCGAGAGAGCGCGGACCGTCGGGTGCTCCGCGCGCCAAACACCGTGTACCTG 180
QY 181 CAGGTGGTGGCAGCGGTAGCGGACTCGGCGCGCGGCTCTACGTCTTCTCCGAGTTC 240
Db 181 CAGGTGGTGGCAGCGGTAGCGGACTCGGCGCGCGGCTCTACGTCTTCTCCGAGTTC 240
```

```
QY 241 AACCGGTATCTCTTCAACTGTGGAGAGCGGTTTCAGAGACTCATGCAGGAGCACAAGTTA 300
Db 241 AACCGGTATCTCTTCAACTGTGGAGAGCGGTTTCAGAGACTCATGCAGGAGCACAAGTTA 300
QY 301 AAGGTTGCTCGCTCGGCAACATATTCCTGACACGAATGCACCTGTCTAATGTTGGGGGC 360
Db 301 AAGGTTGTTGGCTGGCAACATATTCCTGACACGAATGCACCTGTCTAATGTTGGGGGC 360
QY 361 TTAAGTGGAAATGATTTTACTTTTAAAGAAACCGGGCTTCAAAGTGTGTAATTTCTGGA 420
Db 361 TTAAGTGGAAATGATTTTACTTTTAAAGAAACCGGGCTTCAAAGTGTGTAATTTCTGGA 420
QY 421 CCTCCACAACTGGAAAAATACCTCGAAGCAATCAAAATATTTCTGCTCCATTCAGAAAG 480
Db 421 CCTCCACAGCTGGAAAAATACCTCGAAGCAATCAAAATATTTCTGCTCCATTCAGAAAG 480
QY 481 ATAGAACTGGCTGTGCGGCGCCCACTCTGCGCCCAAGATACGAGGATGAACCATGACAGTT 540
Db 481 ATAGAACTGGCTGTGCGGCGCCCACTCTGCGCCCAAGATACGAGGATGAACCATGACAGTT 540
QY 541 TACAGATTCGCCATACACAGTGAACAGAGAGGGGAAAGCAACCAATGCGCAGAGTCCA 600
Db 541 TACCAGATTCGCCATACACAGTGAACAGAGAGGGGAAAGCAACCAATGCGCAGAGTCCA 600
QY 601 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGGATCTTCAGACTCCGAGTCGAATGAAAT 660
Db 601 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGGATCTTCAGACTCCGAGTCGAATGAAAT 660
QY 661 GAGCCACACCTTCCACATGGTGTGTAGCCAGAGAGGGGTTCAGGACTCTTCCCTGGTC 720
Db 661 GAGCCACACCTTCCACATGGTGTGTAGCCAGAGAGGGGTTCAGGACTCTTCCCTGGTC 720
QY 721 GTAGCTTTCATCTGTAAAGCTTCACTTAAAGAGAGGAAACTTCTTGGTGTCTCAAGCAAG 780
Db 721 GTAGCTTTCATCTGTAAAGCTTCACTTAAAGAGAGGAAACTTCTTGGTGTCTCAAGCAAG 780
QY 781 GAGATGGGCTCCAGTTGGGACAGCTGCCATCGCTCCCATCATTTGCTGTCTCAAGAC 840
Db 781 GAGATGGGCTCCAGTTGGGACAGCTGCCATCGCTCCCATCATTTGCTGTCTCAAGAC 840
QY 841 GGGAAAAGCATCACTCATGAAGAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTCTCCA 900
Db 841 GGGAAAAGCATCACTCATGAAGAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTCTCCA 900
QY 901 GATCCTGGTGTCTTTTGTGGTGAATGTCAGATGAAAGCTTCAATTCACCCATC 960
Db 901 GATCCTGGTGTCTTTTGTGGTGAATGTCAGATGAAAGCTTCAATTCACCCATC 960
QY 961 TGTGAGAAATGCCACCTTTTCAGAGTACCAAGGAAAGCAGATGCCCGCTTGGTG 1020
Db 961 TGTGAGAAATGCCACCTTTTCAGAGTACCAAGGAAAGCAGATGCCCGCTTGGTG 1020
QY 1021 GTTCACATGGCCCGCAGCATCTGTGCTTTGTGGACAGAGGTACCAAGAGTGGATGGAGG 1080
Db 1021 GTTCACATGGCCCGCAGCATCTGTGCTTTGTGGACAGAGGTACCAAGAGTGGATGGAGG 1080
QY 1081 TTTGGGCTGACACCCAGCACTTGGTCTCTGAATGAGAACTGTGCTCAGTTCACAACTT 1140
Db 1081 TTTGGGCTGACACCCAGCACTTGGTCTCTGAATGAGAACTGTGCTCAGTTCACAACTT 1140
QY 1141 CGCAGCCACAGATTTCAACCCAGCTCAACCTCATCCACCGGACATCTTCCCGCTGCTC 1200
Db 1141 CGCAGCCACAGATTTCAACCCAGCTCAACCTCATCCACCGGACATCTTCCCGCTGCTC 1200
QY 1201 ACCAGTTTCCGCTGTAAAGAGGGGCCCCCACTCAGTGTGCCCTAGTTCAGGGTGA 1260
Db 1201 ACCAGTTTCCGCTGTAAAGAGGGGCCCCCACTCAGTGTGCCCTAGTTCAGGGTGA 1260
QY 1261 TGCCTCTCAAGTACAGCTCCGTCAGAGGAGGTGGCAGAGGATGCCATTACT 1320
Db 1261 TGCCTCTCAAGTACAGCTCCGTCAGAGGAGGTGGCAGAGGATGCCATTACT 1320
```

```
QY 1321 TCGAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCAGCAGAGCGTG 1380
Db 1321 TCGAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCAGCAGAGCTG 1380
QY 1381 CAGGAGTACAGGAGAGTGCAGCAGAGCGCCAGAGCCCGCAGCAGAGAGAAAGAGTCAGTAC 1440
Db 1381 CAGGAGTACAGGAGAGTGCAGCAGAGCGCCAGAGCCCGCAGCAGAGAGAAAGAGTCAGTAC 1440
QY 1441 CCAGAAATCATCTTCTCTTGGAAACAGGGTCTGCCATCCCGATGAAGATTGAAATGTCAGT 1500
Db 1441 CCAGAAATCATCTTCTCTTGGAAACAGGGTCTGCCATCCCGATGAAGATTGAAATGTCAGT 1500
QY 1501 GCCACATCTGTCAATAGACCCCGACAGCTCTCTGCTACTGACTGCTGGTGAAGGCACA 1560
Db 1501 GCCACATCTGTCAATAGACCCCGACAGCTCTCTGCTACTGACTGCTGGTGAAGGCACG 1560
QY 1561 TTTGGGCACTGTGGCGTCAATACAGGAGACAGGTGACAGGGTCTGGGCAACCTGGCT 1620
Db 1561 TTTGGGCACTGTGGCGTCAATACAGGAGACAGGTGACAGGGTCTGGGCAACCTGGCT 1620
QY 1621 GCTGTGTTTGTCTCCACCTGCAGCAGATCACCAACAGGGCTTGCAGATATCTTGTCTG 1680
Db 1621 GCTGTGTTTGTCTCCACCTGCAGCAGATCACCAACAGGGCTTGCAGATATCTTGTCTG 1680
QY 1681 CAGAGAAAGCGCCTTGGCATCTTTGGAAAGCGCTTACCCCTTGTGCTGTGGCT 1740
Db 1681 CAGAGAAAGCGCCTTGGCATCTTTGGAAAGCGCTTACCCCTTGTGCTGTGGCT 1740
QY 1741 CCCAACCACTCAAGCCTGGCTCCAGCAGTACCAACAGGTGCGAGAGTCTCTGCAC 1800
Db 1741 CCCAACCACTCAAGCCTGGCTCCAGCAGTACCAACAGGTGCGAGAGTCTCTGCAC 1800
QY 1801 CACATCAGTATGATTCCTGCCAAATGCCCTTCAGGAAGGGCTGAGATCTCCAGTCTCTCA 1860
Db 1801 CACATCAGTATGATTCCTGCCAAATGCCCTTCAGGAAGGGCTGAGATCTCCAGTCTCTCA 1860
QY 1861 GTGGAAGAGTATGATGCTGCTGTGCGAATGATGATTTGGAAGAGTTTCAGACCTGT 1920
Db 1861 GTGGAAGAGTATGATGCTGCTGTGCGAATGATGATTTGGAAGAGTTTCAGACCTGT 1920
QY 1921 CTGGTGGGCACTGCAAGCATGCTGTTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGAA 1980
Db 1921 CTGGTGGGCACTGCAAGCATGCTGTTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGAA 1980
QY 1981 GTGGTCTATTCCGGGACACCATGCTCGAGGCTCTGGTCCGATGGGAAAGATGCC 2040
Db 1981 GTGGTCTATTCCGGGACACCATGCTCGAGGCTCTGGTCCGATGGGAAAGATGCC 2040
QY 2041 ACCCTCTGATACATGAAGCCACCTTGGAGATGTTGTTGGAGAGAGCAGTGGAAAG 2100
Db 2041 ACCCTCTGATACATGAAGCCACCTTGGAGATGTTGTTGGAGAGAGCAGTGGAAAG 2100
QY 2101 ACACAGACACAGCTCCCAAGCCATCAGGTGGGATCGGATGAACCGGAGTTCATT 2160
Db 2101 ACACAGACACAGCTCCCAAGCCATCAGGTGGGATCGGATGAACCGGAGTTCATT 2160
QY 2161 ATGCTGAACCACTTCAGCCAGCGCTATGCAAGGTCCCTCTTACGCCCCCACTTCAGC 2220
Db 2161 ATGCTGAACCACTTCAGCCAGCGCTATGCAAGGTCCCTCTTACGCCCCCACTTCAGC 2220
QY 2221 GAGAAAGTGGGAGTGGCTTTGACCAATGAAGTTCCTTTGGAGACTTTTCCAAAGT 2280
Db 2221 GAGAAAGTGGGAGTGGCTTTGACCAATGAAGTTCCTTTGGAGACTTTTCCAAAGT 2280
QY 2281 CCCAGAGTATTCCTCCCACTGAAGCCCTGTTTGGCGGCAATCGAGAGATGGAGGAG 2340
Db 2281 CCCAGAGTATTCCTCCCACTGAAGCCCTGTTTGGCGGCAATCGAGAGATGGAGGAG 2340
QY 2341 CGCAGGAGAGCGGAGTGGCGGAGTGGCGGCGGCTCTCTGCTCCAGGAGCTGGCA 2400
Db 2341 CGCAGGAGAGCGGAGTGGCGGAGTGGCGGCGGCTCTCTGCTCCAGGAGCTGGCA 2400
QY 2401 GCGGCGCTGGAGATGGGGAGCTTCAGCAGAGAGCGGCGCCACACAGAGAGGCCACAGGCC 2460
```

```
Db 2401 GCGCGCCTGGAGATGGGAGCCTCAGCAGAAACGGGCCACACAGAGAGGCCACAGGCC 2460
QY 2461 AAGAAGGTCAGAGCCAGTGA 2481
Db 2461 AAGAAGGTCAGAGCCAGTGA 2481

RESULT 8
US-09-988-687-225
; Sequence 225, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,687
; CURRENT FILING DATE: 2001-11-20
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 225
; LENGTH: 2892
; TYPE: DNA
; ORGANISM: Gorilla gorilla
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
US-09-988-687-225

Query Match 98.5%; Score 2442.6; DB 9; Length 2892;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 2457; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 1 ATGTGGGCGCTTTGCTCGCTGCTGCGGTCGCGGCGCGGACGACCATGTCGCGAGGACGC 60
Db 1 ATGTGGGCGCTTTGCTCGCTGCTGCGGTCGCGGCGCGGACGACCATGTCGCGAGGACGC 60
QY 61 ACCATATCGCAGGACACCGCGCGCGCGCGAGCGGCGCGCAAGGACCCGCTGCGGACCTG 120
Db 61 ACCATATCGCAGGACACCGCGCGCGCGCGAGCGGCGCGCAAGGACCCGCTGCGGACCTG 120
QY 121 CGCAGCGCAGAGAGCGCGGACCGTCCGGGTCCTCCGGGCGCCCAACACCGTGTACCTG 180
Db 121 CGCAGCGCAGAGAGCGCGGACCGTCCGGGTCCTCCGGGCGCCCAACACCGTGTACCTG 180
QY 181 CAGGTGGTGGCAGCGGTAGCGGACTCGGCGCGCGGCTCTAGCTCTTCTCCGAGTTC 240
Db 181 CAGGTGGTGGCAGCGGTAGCGGACTCGGCGCGCGGCTCTAGCTCTTCTCCGAGTTC 240
QY 241 AACCGGTATCTTCAACTGTGGAGAGCGGTTACAGACTCATCGAGAGACACAAGTTA 300
Db 241 AACCGGTATCTTCAACTGTGGAGAGCGGTTACAGACTCATCGAGAGACACAAGTTA 300
QY 301 AAGGTGCTCGCTGGACAAACATATTCCTGACACAAATGCACTGCTTAATGTTGGGGC 360
Db 301 AAGGTGTTTCGCTGGACAAACATATTCCTGACACAAATGCACTGCTTAATGTTGGGGC 360
QY 361 TTAAGTGAATGATTCTTCTTAAAGGAAACCGGGCTTCCAAAGTGTCTACTTTCTGGA 420
Db 361 TTAAGTGAATGATTCTTCTTAAAGGAAACCGGGCTTCCAAAGTGTCTACTTTCTGGA 420
QY 421 COTCCAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGGAAGGA 480
```


Db 421 ||||| CCTCAGAGCTGGAATAAATCTCGAGCANTCAAAATATTTCTGGTCCATTGAAAGGA 480
Qy 481 ATAGAACTGGGTGTGCGGCCACCTCTGCCAGAAATACGAGGATGAAACCATGACAGTT 540
Db 481 ATAGAACTGGGTGTGCGGCCACCTCTGCCAGAAATACGAGGATGAAACCATGACAGTT 540
Qy 541 TACCAGATCCCATACACAGTGAACAGAGGAGGGGAAAGCACCACCATGCGAGAGTCCA 600
Db 541 TACCAGATCCCATACACAGTGAACAGAGGAGGGGAAAGCACCACCATGCGAGAGTCCA 600
Qy 601 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCCGAATGAAAT 660
Db 601 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCCGAATGAAAT 660
Qy 661 GAGCCACACCTTCCACATGGTGTAGCCAGAGAGAGAGGGGTGAGGATCTTCCCTGTC 720
Db 661 GAGCCACACCTTCCACATGGTGTAGCCAGAGAGAGAGGGGTGAGGATCTTCCCTGTC 720
Qy 721 GTAGCTTTTCATCTGAAGTTTCACTTAAAGAGAGAAACTTCTGGTCTCAAGCAAG 780
Db 721 GTAGCTTTTCATCTGAAGTTTCACTTAAAGAGAGAAACTTCTGGTCTCAAGCAAG 780
Qy 781 GAGATGGGCTCCCACTGAGTGGGACAGCTGCCATCGCTCCCATCATTTGCTGTCAAGGAC 840
Db 781 GAGATGGGCTCCCACTGAGTGGGACAGCTGCCATCGCTCCCATCATTTGCTGTCAAGGAC 840
Qy 841 GGGAAAGCATCACTATGAAGGAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTCCA 900
Db 841 GGGAAAGCATCACTATGAAGGAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTCCA 900
Qy 901 GATCTGTGTCTCTTTTGTGTGTGTAGATGTCCAGATGAAAGCTTCAATCAACCCATC 960
Db 901 GATCTGTGTCTCTTTTGTGTGTGTAGATGTCCAGATGAAAGCTTCAATCAACCCATC 960
Qy 961 TGTGAATGTCACCTTTTCAGAGGTACCAAGGAAAGGAGATGCCCGTGGCTTGGT 1020
Db 961 TGTGAATGTCACCTTTTCAGAGGTACCAAGGAAAGGAGATGCCCGTGGCTTGGT 1020
Qy 1021 GTTCATGTGGCCCCAGCATCTGTGTTGTGACAGAGGTACCAAGCATGGATGAGAGG 1080
Db 1021 GTTCATGTGGCCCCAGCATCTGTGTTGTGACAGAGGTACCAAGCATGGATGAGAGG 1080
Qy 1081 TTTGGGCTGACACCCAGCATTTGCTGCTGAATGAACTGTGCTCAGTTTCAACACTT 1140
Db 1081 TTTGGGCTGACACCCAGCATTTGCTGCTGAATGAACTGTGCTCAGTTTCAACACTT 1140
Qy 1141 GCGAGCCACAAGATTCAAACCCAGCTCAACCTCAATCCACCGGACATCTTCCCGCTGTC 1200
Db 1141 GCGAGCCACAAGATTCAAACCCAGCTCAACCTCAATCCACCGGACATCTTCCCGCTGTC 1200
Qy 1201 ACCAGTTTCCGCTGTAAGAGAGGGCCCCACCTCAGTGTGCCCATGGTTCAAGGTGAA 1260
Db 1201 ACCAGTTTCCGCTGTAAGAGAGGGCCCCACCTCAGTGTGCCCATGGTTCAAGGTGAA 1260
Qy 1261 TGCCTCTCAAGTACAGCTCCGTCAGGAGGGAGTGGCAGAGGGATGCCATTATTA 1320
Db 1261 TGCCTCTCAAGTACAGCTCCGTCAGGAGGGAGTGGCAGAGGGATGCCATTATTA 1320
Qy 1321 TGCAATCTGAGGAATTCATAGTTGAGGCGTGTGACAGTTCCTCCAGCAGAGCGTG 1380
Db 1321 TGCAATCTGAGGAATTCATAGTTGAGGCGTGTGACAGTTCCTCCAGCAGAGCGTG 1380
Qy 1381 CAGGAGTACAGGAGGTGCGCAGGACGGCCCCAGCCCCAGAGAGAAAGAAAGTCA 1440
Db 1381 CAGGAGTACAGGAGGTGCGCAGGACGGCCCCAGCCCCAGAGAGAAAGAAAGTCA 1440
Qy 1441 CCAGAAATCATCTTCTTGGACAGGGTCTGCCATCCCGATGAAATGTCAGT 1500
Db 1441 CCAGAAATCATCTTCTTGGACAGGGTCTGCCATCCCGATGAAATGTCAGT 1500
Qy 1501 GCCACACTGTCAACATGAAGCCCCACACGCTCTCTGCTACTGGAGTGTGGTGGAGG 1560

Db 1501 GCCACACTGTCAACATAAGCCCCGACACAGCTCTCTGCTACTGGACTGTGGTGAGG 1560
Qy 1561 TTTGGGAGAGTGTCCGCTCATTTACGGAGACAGAGTGAGAGGGTCTTGGGACACCTGGCT 1620
Db 1561 TTTGGGAGAGTGTCCGCTCATTTACGGAGACAGAGTGAGAGGGTCTTGGGACACCTGGCT 1620
Qy 1621 GCTGTGTTGTGTCCTCCACCTGACGAGATCACACACGGGCTTGCACAGTATCTTGTCTG 1680
Db 1621 GCTGTGTTGTGTCCTCCACCTGACGAGATCACACACGGGCTTGCACAGTATCTTGTCTG 1680
Qy 1681 CAGAGAGAAAGCGGCTTTGGCATCTTTGGGAAAGCGGCTTCAACCTTTGCTGGTGGTGC 1740
Db 1681 CAGAGAGAAAGCGGCTTTGGCATCTTTGGGAAAGCGGCTTCAACCTTTGCTGGTGGTGC 1740
Qy 1741 CCACACAGCTCAAGGCTTGCTCCAGCATCACACACAGTCCAGAGGAGTCTTGAC 1800
Db 1741 CCACACAGCTCAAGGCTTGCTCCAGCATCACACACAGTCCAGAGGAGTCTTGAC 1800
Qy 1801 CACATCAGTATGATTCCTGCTCAATGCTTTCAGGAAGGGCTGAGATCTCCAGTCTGCA 1860
Db 1801 CACATCAGTATGATTCCTGCTCAATGCTTTCAGGAAGGGCTGAGATCTCCAGTCTGCA 1860
Qy 1861 GTGAAAGATTGATCAGTTCGCTGTTGCGAACATGATTTGGAAGAGTTCAGACCTGT 1920
Db 1861 GTGAAAGATTGATCAGTTCGCTGTTGCGAACATGATTTGGAAGAGTTCAGACCTGT 1920
Qy 1921 CTGTGCGGACATGCAAGCATGCGTTTGGCTGTGCGTGTGTCACACTCTGGCTGGA 1980
Db 1921 CTGTGCGGACATGCAAGCATGCGTTTGGCTGTGCGTGTGTCACACTCTGGCTGGA 1980
Qy 1981 GTGCTCTATTTCCGGGACACCATGCCCTGCGAGCTCTGCTCGGATGGGAAAGATGCC 2040
Db 1981 GTGCTCTATTTCCGGGACACCATGCCCTGCGAGCTCTGCTCGGATGGGAAAGATGCC 2040
Qy 2041 ACCCTCTGATACATGAAGCACCCTTGAAGAGTGGTTTGAAGAGGAGTTCGAAAG 2100
Db 2041 ACCCTCTGATACATGAAGCACCCTTGAAGAGTGGTTTGAAGAGGAGTTCGAAAG 2100
Qy 2101 ACACAGCAAGCTCCCAAGCCATCAGCGTGGGATGCGGATGAACGCGGAGTTCA 2160
Db 2101 ACACAGCAAGCTCCCAAGCCATCAGCGTGGGATGCGGATGAACGCGGAGTTCA 2160
Qy 2161 ATGCTGAACCACTTCAGCCAGCGTATGCCAAGTCCCTCTTTCAGCCCCCACTTCAGC 2220
Db 2161 ATGCTGAACCACTTCAGCCAGCGTATGCCAAGTCCCTCTTTCAGCCCCCACTTCAGC 2220
Qy 2221 GAGAAAGTGGGAGTTGCTTTGACCATGAAGTCTGCTTTGGAGACTTTTCCAACTG 2280
Db 2221 GAGAAAGTGGGAGTTGCTTTGACCATGAAGTCTGCTTTGGAGACTTTTCCAACTG 2280
Qy 2281 CCCAAGCTGATTTCCCACTGAAAGCCCTGTTTGTGCGGACATCGAGGAGATGGAGG 2340
Db 2281 CCCAAGCTGATTTCCCACTGAAAGCCCTGTTTGTGCGGACATCGAGGAGATGGAGG 2340
Qy 2341 CGCAGGAGAGGAGGAGCTGCGGAGTGGCGGCGCTCTGCTCCAGGAGCTGCGCA 2400
Db 2341 CGCAGGAGAGGAGGAGCTGCGGAGTGGCGGCGCTCTGCTCCAGGAGCTGCGCA 2400
Qy 2401 GCGGCTTGGAGGATGGGAGCTTCAGCAAGAGCGGCGCCACACAGAGGAGCCACAGGCC 2460
Db 2401 GCGGCTTGGAGGATGGGAGCTTCAGCAAGAGCGGCGCCACACAGAGGAGCCACAGGCC 2460
Qy 2461 AAGAGGTCAAGCCCCAGTGA 2481
Db 2461 AAGAGGTCAAGCCCCAGTGA 2481

RESULT 9

US-09-988-626-221
; Sequence 221, Application US/09988626
; Publication No. US2003004959A1
; GENERAL INFORMATION:
; APPLICANT: Tavtligian, Sean V.


```
Qy 1798 CACCACATCAGTATGATTCCTGCCAAATGCTTCAGGAAGGGCTGAGATCTCCAGTCT 1857
Db 1759 CACCACGTCAGTATGATTCCTGCCAAATGCTTCAGGAAGGGGAGAGGTCTCCAATACT 1818
Qy 1858 GCAGTGGAAAGATGATCAGTTCGCTGTGTGGCAACATGTGATTTGGAAAGATTTTCAGACC 1917
Db 1819 ACATTTGGAAAGGCTGATAAGCTTGCTGTGTGGAAACATGTGACTTAGAAGATTTTCAGACC 1878
Qy 1918 TGTCTGTGGCGGCTGCTCAAGATGCTGTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1977
Db 1879 TGCCTGTGTACGCGCTGCTCAAGCATGCTTTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1938
Qy 1978 AAAGTGTGCTATTTCGGGGGACACCATGCTCGGAGGCTCTGTGCTGCTGTGCTGTGCTGTG 2037
Db 1939 AAAGTGTGCTATTTCGGGGGATACCATGCTCGGAGGCTCTGTGCTGCTGTGCTGTGCTGTG 1998
Qy 2038 GCCACCTCTCTGATACATGAAGCCACCTCGAAGATGCTTTGGAAAGAGGACGATGGAA 2097
Db 1999 GCCACCTCTCTGATACATGAAGCCACTCTGGAGGATCNCCTTGGAAAGAGGACGATAG 2058
Qy 2098 AAGACACACAGACAACTCCCAAGCCATCAGCGTGGGGATGCGGATGAACGCGGAGTTC 2157
Db 2059 AGGACACACAGACCACTCCAGGCTATTAAATGTGGGATGCGGATGAATGCGGAGTTC 2118
Qy 2158 ATTATGCTGAACCACTTCAGCGAGGCTATGCCAGTTCCTCTTCAGCCCACTTC 2217
Db 2119 ATCATGCTGAACCACTTCAGTCAGCGGTACGCAAGATCCCTCTTCAGCCCTGACTTC 2178
Qy 2218 AGCAGAAAGTGGGAGTTCCTTTGACCACATGAAGTCTGCTTTGGAGACTTTTCCAAACA 2277
Db 2179 AACGAAAGTGGATCGCTTTGACCACATGAAGTCTGNTTTGGAGACTTCCCGACA 2238
Qy 2278 ATGCCAAAGCTGATTCCTCCCACTGAAGCCCTGTTTGTGCGCATCGAGAGATGGAG 2337
Db 2239 GTGCCAAAGCTGATTCCTCCCACTGAAGCCCTGTTTGTGCGAGTGACATTTGAAGAGATGGT 2298
Qy 2338 GAGCGCAGGAGAGCGGAGCTGCGGAGTTCGCGGCGGCTCTCTTCAGGAGCTG 2397
Db 2299 GAACGCGGAGAGAGGAGCTACGGCTGTGTGCGAGCAGCCCTCTGACC---CAGCAG 2355
Qy 2398 GCAGCGGCTCTGGAGTGGGAGCTTCAGCAGAAAGCGGGCCCAACACAGAGGAGCCACA 2456
Db 2356 GCAGACAGCCCGAGGACAGAGAACCCCAACAGAAAGCGGGCCCAACACAGATGAACCA 2414

RESULT 10
US-09-988-687-221
; Sequence 221, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,687
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 221
; LENGTH: 2470
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
```

```
; NAME/KEY: CDS
; LOCATION: (1)..(2466)
US-09-988-687-221

Query Match 66.3%; Score 1645.6; DB 9; Length 2470;
Best Local Similarity 81.6%; Pred. No. 0;
Matches 1958; Conservative 0; Mismatches 417; Indels 24; Gaps 4;

Qy 58 CGACCATATTCGAGGACCCCGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 117
Db 40 CGCACCATGTCGAGGGTTCGGCTCGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 99
Qy 118 CTGGCAGCGAGAGAGCGCGGACCGTCCGGGTGCTCCGGCGCGCGCGCGCGCGCGCGCG 177
Db 100 CTGCGTACGCGGAGAGCGCGGCGC-----GGGTCCCGGGGCGCGCGCGCGCGCGCG 153
Qy 178 CTGAGGTGTGGCGGAGCGCGGCTAGCCGGGACTCTCGGCGCGCGCGCGCGCGCGCGCG 237
Db 154 CTGAGGTGTGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 213
Qy 238 TTCAACCGGTATCTCTTCAACTGTGGAGAGGCGTTTCAGAGACTCATGAGGAGGACACA 297
Db 214 TACAACAGGTACCTTTTAACTCGGAGAGGCGTCCCAACGACTTATGCAAGGACACA 273
Qy 298 TTAAGGTGCTCGCTGGGACACATATTCCTGACACGAATGCACTGCTTAATGTTGGG 357
Db 274 ACTGAAAGTCTGCTGTGACACATCTTTCTGACTCGGATGCAATGCTCAATGTTGGG 333
Qy 358 GGCTTAAAGTGAATGATTCTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCT 417
Db 334 GGGTGTGTGGAATGATTTTAACTTTAAAGGAAACCGGGCTTCCCAATGTTCTCTCT 393
Qy 418 GGACCTCCCAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGTCTCAATGAAA 477
Db 394 GGACCCACACAGCTGGAGAAATATCTAGAAGCAATCAAAATATTTTCTGTCTCAATGAAA 453
Qy 478 GGAATAGAACTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 537
Db 454 GGAATAGAACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 513
Qy 538 GTTTTACAGATCCCCATACACAGTGAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 597
Db 514 GTTTTACAGATCCCCATACACAGTGAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 573
Qy 598 CGAGAAAGGCTCTCAGCAGGCTCAGTCCAGAGGATCTTCAGACTCCGAGTCCGAGTGA 657
Db 574 CCCAGAACATCTCCCAACAGGCTCAGTCCCAACAGCTCATCGGACTCTGGATCAGCTGAA 633
Qy 658 AATGAGCCACACTTCCACATGTTGTAGCCAGAGAGAGGAGGAGGAGGAGGAGGAGGAG 717
Db 634 AATGGGC-----AGTGCCCAACAGGAAAGCATGGGCGAGGAG-CTCTCTTA 678
Qy 718 GTCTAGCTTTTCACTCTTAAGCTTCACTTAAAGAGAGGAGGAGGAGGAGGAGGAGGAG 777
Db 679 GTGTAAGCTTTTGTCTGCAAGCTTCACTTGAAGAGAGGAGGAGGAGGAGGAGGAGGAG 738
Qy 778 AAGAGATGGGCTCCAGTGGGACAGCTGCGCATTCCTCCATCATCTGTCTGTCTCAAG 837
Db 739 AAGAGCTGGGCTTCTCTGTGGGACGCGCGCATTTGACCCCATCATTTGTCTGTCTCAAG 798
Qy 838 GACGGGAAAGCATCACTCATGAAGGAGAGAGATTTTGGCTGAAGAGCTGTGTACTCT 897
Db 799 GACGGGAAAGATCACTTACGAAAGAGAGAGATTTGCTGTGAGAGCTTTGTACACCC 858
Qy 898 CCAGATCCTGTGTCTGCTTTTGTGGTGTGAATGTCCAGATGAAAGCTTCAATCAACCC 957
Db 859 CCAGATCCTGTGTCTGCTTTTCACTGTGTAGAGTCTCTGATGAAGGATTCATCTGTGCC 918
Qy 958 ATCTGTGAGATGCCACTTTTCAGAGGTACCAAGGAGGAGGAGGAGGAGGAGGAGGAGG 1017
Db 919 ATCTGTGAGAAACGACCTTTTAAAGGTACCGGACAGAGCTGATGACCTGTGTGGCGCT 978
Qy 1018 GTGTTTCACTATGGCCCGCAGCATCTGTCTTGTGGACAGAGGTACCGAGCTGTGATGGAG 1077
```

```
Db 979 GTGGTCCATAGCCCCAGAACTCTGACTCATCGACAGCAGATACACAGCAGTGGATGGAG 1038
Qy 1078 AGTTTGGGCTTGACACCCAGCAGCTTGGTCTCTGAATGAGAATGTGCTCAGTTCACAAAC 1137
Db 1039 AGTTTGGGCTTGACACACAGCAGCAGCTGATTCGTAATGAGAAATGGCCCTCGTCCACAAAC 1098
Qy 1138 CTTCGGAGCACAAGATTCAAACCCAGCTCAACCTCATCCACCGGACATCTTCCCCCTG 1197
Db 1099 CTGGCAGGCCAAGATTCAGACCCAGCTCAGCTCATCCACCTGACATCTTCCCCAG 1158
Qy 1198 CTCACAGTTTCGGCTGTAAGAGGAGGCGCCACCCCTCAGTGTGCCATGTTTCAGGGT 1257
Db 1159 CTTACAGCTTCTATAGTAAGAGGAGGGTCCACCCTCAGCGTGCCAAACAGTTTCGGGT 1218
Qy 1258 GAATGCTCTCTCAAGTACACAGCTCCGTCCAGGAGGAGTGGCAGAGGATGCCATTATT 1317
Db 1219 GAATGCTCTCTCAAGTATTTCAGTCCGCCCCAAGAGAGAGTGGCAGAGGATACCCACATC 1278
Qy 1318 ACTTGCAATCTTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCACTTCAGCAGAGC 1377
Db 1279 GACTGCAATACTGATGAATTCATAGCTGAGGCGCTTGGAGCTCCCGAGTTTCCAGGAGT 1338
Qy 1378 GTGCAGGAGTACAGGAGGAGTGGCAGGAGCGCCAGCGCCACGACGAGAGAAAGAGTCA 1437
Db 1339 GTGGAGGATTCGGAGAAACGTGCGAGGAAACCCAGCCCCAGCAGAGAGAAAGAGCCAG 1398
Qy 1438 TACCAGAAATCATCTTCTTGGAAACAGGGTCTGCCATCCCGATGAAGATTCGAATGTC 1497
Db 1399 TATCTCGAAATGTTCTTCTGGTACGGGTCTGCCATCCCAATGAGATCCGAATGTC 1458
Qy 1498 AGTGCCACTTGTCAACATAGACCCCGACAGCTCTCTCTACTGACTGTGGTAGGGC 1557
Db 1459 AGTTCCACACTCGTCAACCTTAAGCCCTGACAAAGTCAGTCTCTGGATGTGGAGAGGC 1518
Qy 1558 ACATTGGGACGTGGCGCTCATACGAGAGCAGGTGGAGAGGTCCTGGGACCCCTG 1617
Db 1519 ACTTTGGGACGTGGCGCTCATACGAGACAGCAATAGCCGAGTCTTATGACGCCCTC 1578
Qy 1618 GCTGCTGTGTTGTGTCACCTGACGCGAGATCACCAACGGGCTTGCCAAAGTATCTTG 1677
Db 1579 ACGGCTGTGTTGTGTCACCTGACGCGGACCCACCAACGGGCTTGCTGAATATCTG 1638
Qy 1678 CTGCAGAGAGACGGCCTTGATCTTTGGAAAGCCGCTTCAACCTTTGCTGTGGT 1737
Db 1639 CTGCAGAGAGAGTGGCTTGGCATCTCTGGGAAACCCCTTCAGCCCTTGTGTGGTG 1698
Qy 1738 GCCCCCAACGAGCTCAAGCCTGGCTCCAGCAGTACCAACAGTGGCAGGAGTCTG 1797
Db 1699 GCTCCTACCCAGCTCAGGGCCTGGCTGCGAGCAGTATCAACACCACTGCCCAGGAGTCTG 1758
Qy 1798 CACCACATCAGTATGATTCTGCAAAATGCCCTTCAGGAAGGGCTGAGATCTCCAGTCT 1857
Db 1759 CACCAGCTCAGTATGATTCTGCAAAATGCCCTTCAGGAAGGGCAGAGTCTCCAATACT 1818
Qy 1858 GCAGTGGAAAGATGATCAGTTCGTGTTGGAAACATGTGATTTGGAAGAGTTTCAGACC 1917
Db 1819 ACATTGGAAAGCTGATAAGCTTGCTGTTGGAAACATGTGATTTAGAAATTTTCAGACC 1878
Qy 1918 TGTCTGTTGCGGCACTGCAAGCATGGTTGGCTTGCGCTGCGCTGGCAGACACTCTGGCTGG 1977
Db 1879 TGCTGGTACGGCAGCTGCAAGCATGGTTTGGCTGTGCACTGTGATTCATCTGGCTGG 1938
Qy 1978 AAAGTGGTCTATTTCGGGGACACCATGCCCTTCGAGGCTCTGGTCCGATGGGAAAGAT 2037
Db 1939 AAAGTGGTCTACTTCGGGGATACCATGCCCTGTGAGGCTCTGGTCCAGATGGGAAAGAT 1998
Qy 2038 GCCACCTCTGATACATGAAGCCACCTCTGGAGATGGTTTGGAGAGGAGAGCAGTGGAA 2097
Db 1999 GCCACCTCTGATACATGAAGCCACTCTGGAGGATCNCCTTGGAGAGGAGAGCAGTAGAG 2058
Qy 2098 AAGACACAGACACAGTCCCAAGCCATCAGGTGGGAGTCCGAGTGAACCGGAGTTC 2157
```

```
Db 2059 AGGACACACAGACACACCTCCAGGCTATTAAATGTGGGATGCGGATGAATCGGAGTTC 2118
Qy 2158 ATTATGCTGAACCACTTCAGCCAGCGCTATGCCAAGGTCCCCCTCTTCAGCCCCAACTTC 2217
Db 2119 ATCATGCTGAACCACTTCAGTCAGCGGTACGNAAGATCCCCCTTTTCAGCCCTGACTTC 2178
Qy 2218 AGCGAAGAGTGGGAGTTCCTTTGACCACATGAAGGTCTGCTTTGGAGACTTTTCCAACA 2277
Db 2179 AACGAGAAGTGGCATCGCTTTGACCACATGAAGGTCTGNTTTGGAGACTTCCCGACA 2238
Qy 2278 ATGCCCAAGCTGATTCCCCCACTGAAGCCCTGTTTGTGGCGACATCGAGGAGATGGAG 2337
Db 2239 GTGCCCAAGCTGATTCCCCCACTGAAGCCCTGTTTGCAGGTGACATTTGAAGAGATGGTG 2298
Qy 2338 GAGCCAGGAGAAAGCGGAGCTGCGCAGGTGCGGGCGGCCCTCTCTGTCCAGGAGCTG 2397
Db 2299 GAACGACGGAGAGAGGAGCTACGGCTGTGTGCGAGCAGCCCTCCTGACC--CAGCAG 2355
Qy 2398 GCAGCGGCTTGAGGATGGGAGGCTCAGCAGAAAGCGGGCCACACAGAGAGGCCACA 2456
Db 2356 GCAGACAGCCAGAGGACAGAGAACCCCAAGAGAGCGGGCCACACAGATGAACCACA 2414

RESULT 11
US-09-833-381-2039
; Sequence 2039, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; PRIORITY FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2039
; LENGTH: 783
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(783)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-2039
```

```
Query Match 29.6%; Score 734.8; DB 10; Length 783;
Best Local Similarity 98.3%; Pred. No. 2e-220;
Matches 772; Conservative 0; Mismatches 10; Indels 3; Gaps 3;
```

```
Qy 91 CGGCGCGCAAGGACCCGCTGCGGCACCTGCGCAGAGAGAGCGCGGACCGTCGGGG 150
Db 1 CGGCGCGCAAGGACCCGCTGCGGCACCTGCGCAGAGAGAGCGCGGACCGTCGGGG 60
Qy 151 TGCTCCGCGCGCCCAACACACCGTGTACCTGCAAGTGTGGCAGCGGTAGCGGAGCTCG 210
Db 61 TGCTCCGCGCGCCCAACACACCGTGTACCTGCAAGTGTGGCAGCGGTAG-CCGCGACTCG 119
Qy 211 GCGCGCGCGCTCTAGCTTCTCCGAGTTCAACCGGTATCTTCAACTGTGGAGAGGC 270
Db 120 GCGCGCGCGCTCTAGCTTCTCCGAGTTCAACCGGTATCTTCAACTGTGGAGAGGC 179
Qy 271 GTTCAGACACTCATGCAAGGACACAGTAAAGGTGCTCGCTCGGACACATATTCCTG 330
Db 180 GTTCAGACACTCATGCAAGGACACAGTAAAGGTGCTCGCTCGGACACATATTCCTG 239
Qy 331 ACAGGAATGCAGTGTCTAATGTTGGGGCTTAAAGTGAATGATTTACTTTAAAGGAA 390
Db 240 ACAGGAATGCAGTGTCTAATGTTGGGGCTTAAAGTGAATGATTTACTTTAAAGGAA 299
Qy 391 ACCGGCTTCCAAAGTGTGTTCTGGAATCTCAGACTCCCAACTGGAAAAATACTCTGAAGCA 450
```

Db 300 ACCGGGCTCCAAAGTGTTACTTTCTGACCTCCACAACTGGAATAATACCTCGAAGCA 359
Qy 451 ATCAAAATATTTCTGGTCCATTGAAGGAATAGAACTGGCTGTGGGCCCCACACTCTGCC 510
Db 360 ATCAAAATATTTCTGGTCCATTGAAGGAATAGAACTGGCTGTGGGCCCCACACTCTGCC 419
Qy 511 CCAGAAATACGAGGATGAACCAATGACAGTTTACCAAGATCCCCATACACAGTGAACAGAGG 570
Db 420 CCAGAAATACGAGGATGAACCAATGACAGTTTACCAAGATCCCCATACACAGTGAACAGAGG 479
Qy 571 AGGGGAAGCAACCAACCATGGCAGAGTCCAGAAAGGCTCTCAGCAGGCTCAGTCCAGAG 630
Db 480 AGGGGAAGCAACCAACCATGGCAGAGTCCAGAAAGGCTCTCAGCAGGCTCAGTCCAGAG 539
Qy 631 CGATCTTCAGACTCCGAG-TCGAATGAATAAGAGCAACCTTCCACATGGTGTAGGCA 689
Db 540 CGATCTTCAGACTCCGAGTTCGAATGAATAAGAGCAACCTTCCACATGGTGTAGGCA 599
Qy 690 GAGAAGAGGGTCAGGACTCTTCCTCGTGTAGCTTTCATCTGTAAGCTTCACTTAAA 749
Db 600 GAGAAGAGGGTCAGGACTCTTCCTCGTGTAGCTTTCATCTGTAAGCTTCACTTAAA 659
Qy 750 GAGAGAAACTTCTTGGTCTCAAGCAAGAGAGATGGGCTCCAGTTGGGACAGCTGC 809
Db 660 GAGAGAAACTTCTTGGTCTCAAGCAAGAGAGATGGGCTCCCAATTTGGGAACTGC 718
Qy 810 CATCCTCCATCTGCTGCTCAAGACGGGAAAAGCATCATCTGAAGGAAGAGA 869
Db 719 CATCCTCCATCTGCTGCTCAAGACGNGNAACCAACCACTTAAAGGAAAGA 778
Qy 870 GATTT 874
Db 779 GATTT 783

RESULT 12
US-09-833-381-2038
; Sequence 2038, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: NO. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2038
; LENGTH: 536
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-381-2038

Query Match 19.0%; Score 470.4; DB 10; Length 536;
Best Local Similarity 99.8%; Pred. No. 2.4e-137;
Matches 471; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 91 CGGCCGCGCAAGGACCCGCTGGCGACCTGCGCAGCGGAGAGAGCGCGGACCGTCGGGG 150
Db 1 CGGCCGCGCAAGGACCCGCTGGCGACCTGCGCAGCGGAGAGAGCGCGGACCGTCGGGG 60
Qy 151 TGTCTCGGCGGCGCCAAACACCGTGACTGCGAGGTGGTGGCAGCGGGTAGCGGGACTCG 210
Db 61 TGTCTCGGCGGCGCCAAACACCGTGACTGCGAGGTGGTGGCAGCGGGTAGCGGGACTCG 120
Qy 211 GCGCGCGGCTCAGCTCTTCTCCAGTTCACCGGTATCTTCAACTGTGGAGAGGC 270
Db 121 GCGCGCGGCTCAGCTCTTCTCCAGTTCACCGGTATCTTCAACTGTGGAGAGGC 180
Qy 271 GTTCAGAGACTCATCGAGGAGCAAGTTAAAGGTTGTCGCTCGGCAACATATTCCTG 330

Db 181 GTTCAGAGACTCATGCAGGAGCAAGTTAAAGGTTGCTCGCCTGGACAACATATTCCTG 240
Qy 331 ACAGAAATGCACTGGTCTAATGTTGGGGCTTAAGTGAATGATTTCTTACATTTAAAGAA 390
Db 241 ACAGAAATGCACTGGTCTAATGTTGGGGCTTAAGTGAATGATTTCTTACATTTAAAGAA 300
Qy 391 ACCGGGCTTCCAAAGTGTGTACTTTCTGGACCTCCACAACTGGAAATAATACCTCGAAGCA 450
Db 301 ACCGGGCTTCCAAAGTGTGTACTTTCTGGACCTCCACAACTGGAAATAATACCTCGAAGCA 360
Qy 451 ATCAAAATATTTCTGGTCCATTGAAGGAATAGAACTGGCTGTGGGCCCCACACTCTGCC 510
Db 361 ATCAAAATATTTCTGGTCCATTGAAGGAATAGAACTGGCTGTGGGCCCCACACTCTGCC 420
Qy 511 CCAGAAATACGAGGATGAACCAATGACAGTTTACAGATCCCATACACAGTG 562
Db 421 CCAGAAATACGAGGATGAACCAATGACAGTTTACAGATCCCATACACAGTG 472

RESULT 13
US-09-918-995-8996
; Sequence 8996, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8996
; LENGTH: 554
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(554)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-8996

Query Match 17.4%; Score 432.8; DB 9; Length 554;
Best Local Similarity 91.3%; Pred. No. 1.6e-125;
Matches 502; Conservative 0; Mismatches 44; Indels 4; Gaps 4;

Qy 232 TCCGAGTTCAACCGGTATCTTCAACTGTGGAGAGCGGTCAGAGACTCATCGAGAG 291
Db 7 TCACTGCTCGANGGGGATCGAATTTCTCTGGGAGAGCGGATCAGACATATCGAGAG 66
Qy 292 CACAAGTTAAAGGTTGCTCGCTGGCAACATATTTCTGCACCAATGCACTGGTCTAAT 351
Db 67 CACAAGTTAAAGGTTGCTCGCTGGCAACATATTTCTGCACCAATGCACTGGTCTAAT 126
Qy 352 GTTGGGGCTTAAAGTGAATGATTTCTTTAAAGAAACCGGGCTTCAAAAGTGTA 411
Db 127 GTTGGGGCTTAAAGTGAATGATTTCTTTAAAGAAACCGGGCTTCAAAAGTGTA 186
Qy 412 CTTTCTCGACTCCACAACCTGGAAAATACTCGAAGCAATCAAAATATTTCTGGTCCA 471
Db 187 CTTTCTCGACTCCACAACCTGGAAAATACTCGAAGCAATCAAAATATTTCTGGTCCA 246
Qy 472 TTGAAGGAATAGAACTGGCTGTGGGCCCCCACTCTGCCAGAAATACGAGGATGAAC 531
Db 247 TTGAAGGAATAGAACTGGCTGTGGGCCCCCACTCTGCCAGAAATACGAGGATGAAC 306
Qy 532 ATGACAGTTTACAGATCCCATACACAGTGAACAGAGAGGGGAAAGCACCACCATGG 591
Db 307 ATGACAGTTTACAGATCCCATACACAGTGAACAGAGAGGGGAAAGCACCACCATGG 366
Qy 592 CAGAGTCCAGAAAGGCTCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCG 651

Db 367 CAGAGTCCAGAAAGGCTCTCAGCAGGCTCAGTCCAGAGCATCTTCAGACTCCGAGTCG 426
Qy 652 AATGAAATGAGCCACACCTTCACATGGTGTAGCCAGAGAGAGGGGTGACGGACTCT 711
Db 427 AATGAAATGAGCCACACCTTCACATGGTGTAGCCAGAGAGAGGGGTGACGGACTCT 484
Qy 712 TCCTGGTGGTGTAGCTTTCATCTGTAGCTTCACTTAAAGAGAGAGAACTTCTTGGTGCTC 771
Db 485 TCCT-GTGGTGGTGTAGCTTTCATCTGTAGCTTCACTTAAAGAGAGAGAACTTCTTGGTGCTC 542
Qy 772 AAAGCAAGG 781
Db 543 AAAGATGG 552

RESULT 14
US-09-988-626-210
; Sequence 210, Application US/09988626
; Publication No. US2003004959A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT FILING DATE: 2001-11-20
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 210
; LENGTH: 350
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (51)..(293)
US-09-988-626-210

Query Match 10.0%; Score 247.4; DB 9; Length 350;
Best Local Similarity 97.7%; Pred. No. 2.7e-67;
Matches 251; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 1 ATGTGGCGCTTTGCTCGTGTGCTGCGTCCGGCGGAGCGGACCATGTGCGAGGAGCG 60
Db 51 ATGTGGCGCTTTGCTCGTGTGCTGCGTCCGGCGGAGCGGACCATGTGCGAGGAGCG 110
Qy 61 ACCATATCGAGGACACCGCCCGCGGAGCGGCGGAGGACCGCTGCGGACCTG 120
Db 111 ACCATATCGAGGACACCGCCCGCGGAGCGGCGGAGGACCGCTGCGGACCTG 170
Qy 121 CGCAGCGAGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 180
Db 171 CGCAGCGAGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 230
Qy 181 CAGGTGGTGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 240
Db 231 CAGGTGGTGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 290
Qy 241 AACCGGTATCTTTCAA 257
Db 291 AACCGGTAGTCAACGA 307

Search completed: May 18, 2003, 00:05:36
Job time : 322.542 secs

RESULT 15
US-09-988-687-210
; Sequence 210, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT FILING DATE: 2001-11-20
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 210
; LENGTH: 350
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (51)..(293)
US-09-988-687-210

Query Match 10.0%; Score 247.4; DB 9; Length 350;
Best Local Similarity 97.7%; Pred. No. 2.7e-67;
Matches 251; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 1 ATGTGGCGCTTTGCTCGTGTGCTGCGTCCGGCGGAGCGGACCATGTGCGAGGAGCG 60
Db 51 ATGTGGCGCTTTGCTCGTGTGCTGCGTCCGGCGGAGCGGACCATGTGCGAGGAGCG 110
Qy 61 ACCATATCGAGGACACCGCCCGCGGAGCGGCGGAGGACCGCTGCGGACCTG 120
Db 111 ACCATATCGAGGACACCGCCCGCGGAGCGGCGGAGGACCGCTGCGGACCTG 170
Qy 121 CGCAGCGAGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 180
Db 171 CGCAGCGAGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 230
Qy 181 CAGGTGGTGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 240
Db 231 CAGGTGGTGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 290
Qy 241 AACCGGTATCTTTCAA 257
Db 291 AACCGGTAGTCAACGA 307

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 15:56:54 ; Search time 3376.55 Seconds
(without alignments)
11900.029 Million cell updates/sec

Title: US-09-434-382-1
Perfect score: 2481
Sequence: 1 atgtggggcgtttctgcgt.....agaaggtcagagcccaagtga 2481

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_nam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1661	66.9	2875	11 BC023134	BC023134 Mus muscu
2	917.4	37.0	1510	11 AK004136	AK004136 Mus muscu
3	908.8	36.6	1030	14 B0054697	B0054697 AGENCOURT
4	876	35.3	930	14 B0680295	B0680295 AGENCOURT
5	872.2	35.2	918	14 BQ706437	BQ706437 AGENCOURT
6	868.8	35.0	1012	14 BQ066524	BQ066524 AGENCOURT

7	854.4	34.4	1067	14	BM927353	BM927353 AGENCOURT
8	831.2	33.5	915	14	BQ890505	BQ890505 AGENCOURT
9	828.8	33.4	1042	14	BQ057402	BQ057402 AGENCOURT
10	821.6	33.1	978	14	BQ064858	BQ064858 AGENCOURT
11	820.8	33.1	970	14	BQ071202	BQ071202 AGENCOURT
12	801.8	32.3	1009	14	BQ059218	BQ059218 AGENCOURT
13	800	32.2	983	14	BQ069865	BQ069865 AGENCOURT
14	799.4	32.2	876	14	BQ068069	BQ068069 AGENCOURT
15	799	32.2	928	14	BQ679122	BQ679122 AGENCOURT
16	793.2	32.0	881	14	BQ084266	BQ084266 AGENCOURT
17	769.6	31.0	1002	14	BM928586	BM928586 AGENCOURT
18	767	30.9	898	14	BQ923185	BQ923185 AGENCOURT
19	765.4	30.9	791	12	BG822529	BG822529 602725613
20	755	30.4	1065	13	B1825701	B1825701 603072685
21	751.2	30.3	930	14	BQ962958	BQ962958 AGENCOURT
22	736.2	29.7	966	12	BG674022	BG674022 602620010
23	731	29.5	1042	14	BM927272	BM927272 AGENCOURT
24	727.8	29.3	980	14	BQ677668	BQ677668 AGENCOURT
25	724.4	29.2	883	12	BG746363	BG746363 602703625
26	721.2	29.1	960	12	BF797306	BF797306 602256926
c	720.6	29.0	776	13	B1819565	B1819565 603036794
27	720.6	29.0	874	14	BQ674529	BQ674529 AGENCOURT
28	717.4	28.9	874	14	BQ674529	BQ674529 AGENCOURT
29	715.4	28.8	928	12	BG675078	BG675078 602621324
30	709.6	28.6	761	12	BE759820	BE759820 601590856
31	702	28.3	827	12	BG396395	BG396395 602459323
32	699.2	28.2	865	13	B1859833	B1859833 603386144
33	698.2	28.1	762	9	AU124671	AU124671 AU124671
c	692.4	27.9	708	14	BQ109476	BQ109476 imagec 7
34	689.4	27.8	920	12	BG386090	BG386090 602455264
35	687.4	27.7	860	14	BQ642764	BQ642764 AGENCOURT
36	687.4	27.7	842	13	B1601088	B1601088 603249579
37	686.8	27.7	712	9	AU126037	AU126037 AU126037
38	685	27.6	1123	14	BM810990	BM810990 AGENCOURT
39	679	27.4	823	10	BE619259	BE619259 601473130
40	677.2	27.3	771	12	BG766403	BG766403 602739146
41	676	27.3	984	12	BE744876	BE744876 601576324
42	676	27.2	676	10	BE260495	BE260495 601150702
43	674.2	27.2	1012	12	BG386348	BG386348 602455550
44	673.8	27.2	1102	14	BQ056556	BQ056556 AGENCOURT
45						

ALIGNMENTS

RESULT 1	BC023134	BC023134	2875 bp	mRNA	linear	HTC 07-AUG-2002
LOCUS	BC023134	Mus musculus, Similar to elac homolog 2 (E. coli), Clone				
DEFINITION	IMAGE:4018447, mRNA.					
ACCESSION	BC023134					
VERSION	BC023134.1	GI:18605648				
KEYWORDS	HTC.					
SOURCE	house mouse.					
ORGANISM	Mus musculus					
REFERENCE	1 (bases 1 to 2875)					
AUTHORS	Strausberg, R.					
TITLE	Direct Submission					
JOURNAL	Submitted (04-FEB-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA					
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov					
COMMENT	Contact: MGC help desk Email: cgapbs-remail.nih.gov Tissue Procurement: Gilbert Smith, Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: http://www.hgsc.bcm.tmc.edu/cdna/					

1786 CAGGAGTCTCTGACACATCATGATGATTCCTGCGCAATGCTTCCAGAAAGGGCTGAG 1845
 1789 CAGGAGATTTCTGCACACCTCAGTATGATTCCTGCGCAATGCTTCCAGAAAGGGCTGAG 1848
 1846 ATCTCCAGTCTCTGACGAGGAGATGATGATGCTGCTGCGACATGCTGATTTGGAA 1905
 1849 GTCTCAATACATACATGGAAGGCTGATGATGCTGCTGGAACATGCTGACTTAGAA 1908
 1906 GAGTTTTCAGACCTGCTGCTGGGGCACTCAAGCATGCGTTTGGCTGTGCGGTGGTGCAC 1965
 1909 GAATTTTCAGACCTGCTGCTGACGCACTCAACATGCTTTTGGCTGCGCATGCTGATCAT 1968
 1966 ACTCTGCTGGAAGTGTCTATTCCGGGGACACCATGCGCTGCGAGCTCTGTGTCGG 2025
 1969 TCATCTGCTGGAAGTGTCTACTCGGGGATACCATGCGCTGTGAGGCTCTGTGTCACG 2028
 2026 ATGGGGAAGATGCCACCTCTCTGATACATGAGCCACCTGGAAGATGTTTGGGAAGAG 2085
 2029 ATGGGGAAGATGCCACCTCTCTGATACATGAGCCACCTGGAAGATGTTTGGGAAGAG 2088
 2086 GAAGCAGTGGAAAGACACACAGCAACAGCTCCCAAGCCATCAGCTGGGGATGCGGATG 2145
 2089 GAAGCAGTAGAGAGACACACAGCACCACTCCAGGCTATTAACTGGGGATGCGGATG 2148
 2146 AAGCGGAGTTCATATGCTGGAACCACTTCAGCAGCGGTATGCCAAGTCCCTCTTC 2205
 2149 AATCGCGAGTTCATATGCTGGAACCACTTCAGCAGCGGTATGCCAAGTCCCTCTTC 2208
 2206 AGCCCAACTTCAGCAGAGAGTGGAGTTCCTTTGACCAATGAGCTCTGCTTTGA 2265
 2209 AGCCCTGACTTCAACGAGAAAGTGGCATCGCTTTGACCAATGAGCTCTGCTTTGA 2268
 2266 GACTTTCCCAATGCTCCAGCTGATTTCCCACTGAAAGCCCTGTTTGTGCGGACATC 2325
 2269 GACTTTCCCACTGCTCCAGCTGATTTCCCACTGAAAGCCCTGTTTGTGCGATGACAT 2328
 2326 GAGGAGATGAGAGCGCAGGAGAGAGCGGAGCTCGCGAGGTGCGGCGGCTCTCTG 2385
 2329 GAAGAGATGCTGGAACGAGGAGAGAGGAGCTACGCTGCTGCGAGCGCTCTCTG 2388
 2386 TCCAGGAGCTGCGAGCGCTGAGGAGTGGAGCTTCCAGAGCTTCCAGAGCGGCGCCACAC 2445
 2389 ACC---CAGCAGGACAGACGCCAGAGGACAGAGAACCCCAAGAGCGGCGCCACAC 2445
 2446 GAGGA-----GCCACAGGCGCAAGAGTCCAGAGCCAGTGA 2481
 2446 GATGAACCCACAGCCACAGACAGACAGAGAGTCCAGAGCCAGTGA 2490

RESULT 2
 AK004136
 LOCUS
 DEFINITION Mus musculus 18 days embryo whole body cDNA, RIKEN full-length
 enriched library, clone:110037C17:DNA segment, Chr 11, Wayne State
 University 80, expressed, full insert sequence.
 ACCESSION AK004136
 VERSION AK004136.1 GI:12835201
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (strain:C57BL/6J) 18 days embryo cDNA to mRNA,
 clone:lib:RIKEN full-length enriched mouse cDNA library
 clone:110037C17.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 REFERENCE
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 2
 REFERENCE
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
 Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
 Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yananaka, I.,
 Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
 Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
 Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
 Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G.,
 Quackenbush, J., Schriml, L., Stabili, F., Suzuki, R., Tomita, M.,
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
 Baldarelli, R., Barch, G., Blake, J., Boffelli, D., Bojunga, N.,
 Carninci, P., de Bonaldo, M., Brownstein, M., Bult, C., Hill, D.,
 Hofmann, M., Hume, D., Kamiya, M., Lee, N. H., Lyons, P.,
 Mazzarelli, J., Mombauer, P., Nordone, P.,
 Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
 Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,
 Toyooka, K., Wang, K. H., Weiss, C., Wittaker, C., Wilming, L.,
 Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kotsuki, S.
 and Hayashizaki, Y.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)
 21085660
 MEDLINE
 PUBMED 11217851
 5 (bases 1 to 1510)
 AUTHORS Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
 Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C.,
 Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T.,
 Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F.,
 Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S.,
 Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,
 Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J.,
 Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
 Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,
 Sobabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
 Tanaka, T., Tejima, Y., Toyota, T., Yamamura, T., Yananaka, I.,
 Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and
 Hayashizaki, Y.
 Direct Submission
 Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail:genome-research.riken.go.jp,
 URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
 Fax:81-45-503-9216)
 Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
 further details.
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues. First strand cDNA was primed with a primer

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1030)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov

Tissue Procurement: Lou Staudt

cDNA Library Preparation: Rubin Laboratory
cDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: LCM2040 row: p column: 17
High quality sequence stop: 647.

FEATURES
source

Location/Qualifiers

1..1030

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5804032"

/clone_lib="NIH MGC 99"

/tissue_type="lymphoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

BASE COUNT 272 a 266 c 256 g 231 t 5 others

Query Match 36.6%; Score 908.8; DB 14; Length 1030;

Best Local Similarity 98.3%; Pred. No. 1.4e-222;

Matches 939; Conservative 0; Mismatches 13; Indels 3; Gaps 2;

QY 336 AATGCACGTGGTCTAATGTTGGGGGCTTAAGTGAATGATCTTACTTTAAAGGAACCGG 395

Db 1 AATGCACGTGGTCTAATGTTGGGGGCTTAAGTGAATGATCTTACTTTAAAGGAACCGG 60

QY 396 GCTTCCAAAGTGTACTTCTGGACCTCCACAACCTGGAAAAATACCTCGAAGCAATCAA 455

Db 61 GCTTCCAAAGTGTACTTCTGGACCTCCACAACCTGGAAAAATACCTCGAAGCAATCAA 120

QY 456 AATATTTCTGTGTCATTGAAAGGAATAGAACTGGCTGTGGCGGCCCACTCTGCCCCAGA 515

Db 121 AATATTTCTGTGTCATTGAAAGGAATAGAACTGGCTGTGGCGGCCCACTCTGCCCCAGA 180

QY 516 ATACGAGGATGAACCATGACAGTTTACAGATCCCATACAGTGAACAGAGAGGGG 575

Db 181 ATACGAGGATGAACCATGACAGTTTACAGATCCCATACAGTGAACAGAGAGGGG 240

QY 576 AAAGCACCACCATGCGAGGTCGAGAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATC 635

Db 241 AAAGCACCACCATGCGAGGTCGAGAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATC 300

QY 636 TTCAGACTCCGAGTGAATGAAATAGAGCCACACCTTCCACATGGTGTAGCCAGAGAG 695

Db 301 TTCAGACTCCGAGTGAATGAAATAGAGCCACACCTTCCACATGGTGTAGCCAGAGAG 360

QY 696 AGGGGTGACGGGACTCTTCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 755

Db 361 AGGGGTGACGGGACTCTTCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 420

QY 756 AAACCTCTTGGTGCTCAAAAGCAAGAGATGGGCTCCAGTGGGACAGCTGCCATCGC 815

Db 421 AAACCTCTTGGTGCTCAAAAGCAAGAGATGGGCTCCAGTGGGACAGCTGCCATCGC 480

QY 816 TCCCATCATTTGCTGTCAGGACGGGAAAGAGCATCACTCATGAAGGAAGAGATTTT 875
Db 481 TCCCATCATTTGCTGTCAGGACGGGAAAGAGCATCACTCATGAAGGAAGAGATTTT 540
QY 876 GGCTGAAGAGCTGTGTACTCTCTCCAGATCCTGGTGTCTGCTTTTGTGGTGAATGTCC 935
Db 541 GGCTGAAGAGCTGTGTACTCTCTCCAGATCCTGGTGTCTGCTTTTGTGGTGAATGTCC 600
QY 936 AGATGAAGAGCTTCAATCAACCCATCTGTGAGAAATGCCACCTTTTCAGAGGTACAAGGAAA 995
Db 601 AGATGAAGAGCTTCAATCAACCCATCTGTGAGAAATGCCACCTTTTCAGAGGTACAAGGAAA 660
QY 996 GGCAGATCCCCCGTGGCTTGGTGTTCACATGCCGCCAGCATCTGTCTTGTGGACAG 1055
Db 661 GGCAGATCCCCCGTGGCTTGGTGTTCACATGCCGCCAGCATCTGTCTTGTGGACAG 720
QY 1056 CAGGTACACAGAGTGGATGGAGAGGTTTGGGCTTGACACCCAGCATTTGGTCTCTGAATGA 1115
Db 721 CAGGTACACAGAGTGGATGGAGAGGTTTGGGCTTGACACCCAGCATTTGGTCTCTGAATGA 780
QY 1116 GAATGTGCTCAGTTCACAACTTTCGAGGCCACAAAGATTCAAAACCCAGCTCAACCTCAT 1175
Db 781 GAATGTGCTCAGTTCACAACTTTCGAGGCCACAAAGATTCAAAACCCAGCTCAACCTCAT 840
QY 1176 CCACCGGACATCTTCCCTCTCACCAGTTTCCGCTGTAAAGAGGGGCCCCACCCCT 1235
Db 841 CCACCGGACATCTTCCCTCTCACCAGTTTCCGCTGTAAAGAGGGGCCCCACCCCT 900
QY 1236 CAGTGTGCCCAT-GGTTCCAGGGTGA--TGCCTCTCTCAAGTACCAGTCCGTCCTCC 1287
Db 901 CAGTGTGCCCATGGTTTCCAGGGTGAATGTCTCTCTCAAGTACCAGTCCGTCCTCC 955

RESULT 4

LOCUS

BO880295

DEFINITION

5', mRNA sequence.

ACCESSION

BO880295

VERSION

BO880295.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 930)

AUTHORS

NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LCM2533 row: d column: 18

High quality sequence stop: 699.

Location/Qualifiers

1..930

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:6339041"

/clone_lib="NIH_MGC 47"

/tissue_type="neuroblastoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:

EcoRI; cDNA made by oligo-dT priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'

adaptor: GGCACGAG(G). Size-selected >500bp for average

insert size 1.8kb. Library constructed by Ling Hong in


```

Db 309 AGGTTGCTCGCTGACAAATATTCCTGACACGAATGCATGGTCTTAATGTGGGGC 368
Qy 361 TTAAGTGAATGATCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTACTTTCTGGA 420
Db 369 TTAAGTGAATGATCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTACTTTCTGGA 428
Qy 421 CTCCACACTGGAAAAATACCTCGAGCAATCAAAATATTTCTGGTCCATTGAAGGA 480
Db 429 CTCCACAACTGGAAAAATACCTCGAGCAATCAAAATATTTCTGGTCCATTGAAGGA 488
Qy 481 ATAGAACTGGGTGGGGCCCACTCTGCCAGAAATACGAGATGAAACCATGACAGTT 540
Db 489 ATAGAACTGGGTGGGGCCCACTCTGCCAGAAATACGAGATGAAACCATGACAGTT 548
Qy 541 TACCAGATCCCATACACAGTGAACAGAGAGGGGAAAGCACCACCATGGCAGAGTCCA 600
Db 549 TACCAGATCCCATACACAGTGAACAGAGAGGGGAAAGCACCACCATGGCAGAGTCCA 608
Qy 601 GAAAGGCTCTCAGCAGGCTCAGTCAGACAGCGATCTTCAGACTCCAGTCGAATGAAAT 660
Db 609 GAAAGGCTCTCAGCAGGCTCAGTCAGACAGCGATCTTCAGACTCCAGTCGAATGAAAT 668
Qy 661 GAGCCACACTTCCACATGCTGTAGCCAGAGAGAGGGGTGAGGACTCTTCCCTGGTC 720
Db 669 GAGCCACACTTCCACATGCTGTAGCCAGAGAGAGGGGTGAGGACTCTTCCCTGGTC 728
Qy 721 GTAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGG-AAAATCTTCTGGTGTCTCAAAAGCAA 779
Db 729 GTAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGGAAATCTTCTGGTGTCTCAAAAGCAA 788
Qy 780 GGAGATGGGCTCCCGATGGGACAGCTGCCATCGCTCCCATCATGCTGTGTCGAAGA 839
Db 789 GGAGATGGGCTCCCGATGGGACAGCTGCCATCGCTCCCATCATGCTGTGTCGACGA 848
Qy 840 CGGAAAAACATCACTCATGAGGAGAGA-GAGATTTTGGCTGAAGACTGTGTACTCTC 898
Db 849 CGGAAAAACATCACTCATGAGGAGAGAGATTTTGGCTGAAGAGCTGGGGACTCTCTC 908
Qy 899 CAGATCCTG 907
Db 909 CAAATCCTG 917

```

```

RESULT 6
BQ066524
LOCUS AGENCOURT_6869738 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5931173
DEFINITION 1012 bp mRNA linear EST 02-APR-2002
5', mRNA sequence.
ACCESSION BQ066524
VERSION BQ066524.1 GI:19895570
KEYWORDS EST.
SOURCE human.

```

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NIH-MGC <http://mgi.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Lou Staudt

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLCM2112 row: b column: 06

High quality sequence start: 16

High quality sequence stop: 672.

Location/Qualifiers

FEATURES

```

source
1..1012
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5931173"
/clone_lib="NIH_MGC_99"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lymph; Vector: pORF7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dr priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
BASE COUNT 264 a 265 c 260 g 220 t 3 others
ORIGIN
Query Match 35.0%; Score 868.8; DB 14; Length 1012;
Best Local Similarity 96.6%; Pred. No. 2.7e-212;
Matches 950; Conservative 0; Mismatches 25; Indels 8; Gaps 6;
Qy 342 CTGGTCTAATGTGGGGCTTAAGTGAATGATCTTAAAGGAAACCGGGCTTCC 401
Db 30 CTGGTCTAATGTGGGGCTTAAGTGAATGATCTTAAAGGAAACCGGGCTTCC 89
Qy 402 AAAGTGTGTACTTCTGGACCTCCCAACTGGAATAATACCTCGAAGCAATCAAAATATT 461
Db 90 AAAGTGTGTACTTCTGGACCTCCCAACTGGAATAATACCTCGAAGCAATCAAAATATT 149
Qy 462 TTCTGGTCAATGAAAGGAATAGAACTGGCTGTGGGGCCCACTCTGCCCCAGAAATACGA 521
Db 150 TTCTGGTCAATGAAAGGAATAGAACTGGCTGTGGGGCCCACTCTGCCCCAGAAATACGA 209
Qy 522 GGATGAACCATGACAGTTTACAGATCCCCATACACAGTGAACAGAGAGGGGAAAGCA 581
Db 210 GGATGAACCATGACAGTTTACAGATCCCCATACACAGTGAACAGAGAGGGGAAAGCA 269
Qy 582 CCAACCATGGCAGAGTCCAGAAAGGCCTCTCAGCAGGCTCAGTCAGAGCGATCTTCAGA 641
Db 270 CCAACCATGGCAGAGTCCAGAAAGGCCTCTCAGCAGGCTCAGTCAGAGCGATCTTCAGA 329
Qy 642 TTCCGAGTGAATGAAATAGCCACACCTTCCACATGGTGTAGCCAGAGAGAGGGGT 701
Db 330 TTCCGAGTGAATGAAATAGCCACACCTTCCACATGGTGTAGCCAGAGAGAGGGGT 389
Qy 702 CAGGAGCTTCTCCCTGGTGTAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGAACTT 761
Db 390 CAGGAGCTTCTCCCTGGTGTAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGAACTT 449
Qy 762 CTTGGTGTCAAAAGCAAAAGGAGATGGGCTCCAGTTGGGACAGCTGCCATCGCTCCCAT 821
Db 450 CTTGGTGTCAAAAGCAAAAGGAGATGGGCTCCAGTTGGGACAGCTGCCATCGCTCCCAT 509
Qy 822 CATTCGTGCTGTCAGGAGCGGGAAGAGCATCACTCATGAAGGAAGAGAGATTTGGCTGA 881
Db 510 CATTCGTGCTGTCAGGAGCGGGAAGAGCATCACTCATGAAGGAAGAGAGATTTGGCTGA 569
Qy 882 AGAGCTGTGTACTCTCCAGATCCCTGGTGTCTCTTTTGTGGTGTAGATGTCCAGATGA 941
Db 570 AGAGCTGTGTACTCTCCAGATCCCTGGTGTCTCTTTTGTGGTGTAGATGTCCAGATGA 629
Qy 942 AAGCTTTCATTAACCCATCTGTGAGAATGCCACCTTTTCAGAGGTACCAAGGAAAGGCAGA 1001
Db 630 AAGCTTTCATTAACCCATCTGTGAGAATGCCACCTTTTCAGAGGTACCAAGGAAAGGCAGA 689
Qy 1002 TGCCCCCGGT-GGCTTGTGGTTCATCATGGGCCCCAGCATCTCTGTCTTGTGGACAGCAGT 1060
Db 690 TGCCCCCGGTGGGCTTGTGGTTCATCATGGGCCCCAGCATCTCTGTCTTGTGGACAGCAGT 749
Qy 1061 ACCAGAGTGTGATGAGAGGTTTGGGCTGTGACCCAGCAGCTTGTGCTCTGAATGAGAACT 1120

```

```

Db 750 ACCACGATGGATGAGAAATTTGGGCTTGACCCAGCACTTGGTCTGTAATGAGAACT 809
Qy 1121 GTGCCTCAGTTCACAACTTTCGACGCCACAGATTCAAA-CCAGCTCAACTTCATCCAC 1179
Db 810 GTGCCTCAGTTCACAACTTTCGACGCCACAGATTCAAA-CCAGCTCAACTTCATCCAC 869
Qy 1180 CC-GGACATCTTCCCTCGCTCACCAGTTTCGGCT-GTAAGAGGAGGCGCCACCCCTCA 1237
Db 870 CCGGACATCTTCCCTCGCTCACCAGTTTCGGCT-GTAAGAGGAGGCGCCACCCCTCA 929
Qy 1238 GTGTCCCATGG-TTCAGGGTGAATGCTCTCAAGTACCAGCTCCGTCGCCAGGA---GG 1293
Db 930 GGGGCCCCATGTTTCAGGGTGAATGCTCTCAGNTACAGCTCGNTCCAAAGGA 989
Qy 1294 GAGTGGCAGAGGATGCCATTAT 1316
Db 990 AATTGGCAAAGGATGCCATTTT 1012

RESULT 7
BM927353 1067 bp mRNA linear EST 12-MAR-2002
LOCUS AGENCOURT 6704919 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5796765
DEFINITION 5', mRNA sequence.
ACCESSION BM927353
VERSION BM927353.1 GI:19377732
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1067)
NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: sgapbs-remail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2022 row: a column: 22
High quality sequence stop: 650.
Location/Qualifiers
1..1067
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5796765"
/clone_lib="NIH_MGC_99"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
BASE COUNT 277 a 280 c 266 g 237 t 7 others
ORIGIN
Query Match 34.4%; Score 854.4; DB 14; Length 1067;
Best Local Similarity 97.3%; Pred. No. 1.4e-208;
Matches 880; Conservative 0; Mismatches 21; Indels 3; Gaps 1;

Qy 336 AATGCACTGGTCTAATGTTGGGGCTTAAGTGAATGATTCTTAAAGAAACCGG 395
Db 1 AATGCACTGGTCTAATGTTGGGGCTTAAGTGAATGATTCTTAAAGAAACCGG 60

```

```

Qy 396 GCTTCCAAAGCTGTCTATTTCTTGACCTCCACAACTCGAAAAATACCTCGAAGCAATCAA 455
Db 61 GCTTCCAAAGCTGTCTATTTCTTGACCTCCACAACTCGAAAAATACCTCGAAGCAATCAA 120
Qy 456 AATATTTTCTGGTCCATTTGAAAGGAATAGAACTGGCTGTGGGGCCCACTCTGCCCCAGA 515
Db 121 AATATTTTCTGGTCCATTTGAAAGGAATAGAACTGGCTGTGGGGCCCACTCTGCCCCAGA 180
Qy 516 ATACGAGATGAACCATGACAGTTTACCAGATCCCATACACAGTGAACAGAGAGGGG 575
Db 181 ATACGAGATGAACCATGACAGTTTACCAGATCCCATACACAGTGAACAGAGAGGGG 240
Qy 576 AAAGCACCACCATGGCAGATCCAGAAAGCCCTCTCAGCAGGCTCAGTCCAGAGCGATC 635
Db 241 AAAGCACCACCATGGCAGATCCAGAAAGCCCTCTCAGCAGGCTCAGTCCAGAGCGATC 300
Qy 636 TTCAGACTCCGAGTCGAATGAAATGAGCCACACCTTCCACATGGTGTAGCCAGAGAAG 695
Db 301 TTCAGACTCCGAGTCGAATGAAATGAGCCACACCTTCCACATGGTGTAGCCAGAGAAG 360
Qy 696 AGGGTCAGGACCTTCTCCCTGGTCGTAGCTTTTCACTGTGAAGCTTCACTTAAAGAGAGG 755
Db 361 AGGGTCAGGACCTTCTCCCTGGTCGTAGCTTTTCACTGTGAAGCTTCACTTAAAGAGAGG 420
Qy 756 AAATCTTCTGGTGTCAAGCAAGAGGATGGGCTCCAGTTCGGACAGCTGCCATCGC 815
Db 421 AAATCTTCTGGTGTCAAGCAAGAGGATGGGCTCCAGTTCGGACAGCTGCCATCGC 480
Qy 816 TCCCATCATTTCTGCTGTCACAGGACGGGAAAGCATCACTCATGAAGGAAGAGATTTT 875
Db 481 TCCCATCATTTCTGCTGTCACAGGACGGGAAAGCATCACTCATGAAGGAAGAGATTTT 540
Qy 876 GGCTGAAGAGCTGTGTACTCTCCAGATCTCTGGTGTCTCTTTTGTGGTGTAGATGTCC 935
Db 541 GGCTGAAGAGCTGTGTACTCTCCAGATCTCTGGTGTCTCTTTTGTGGTGTAGATGTCC 600
Qy 936 AGATGAAGCTTCAATTCACACCATCTGTGAGAAATGCCACCTTTCAGAGGTACCAAGGAAA 995
Db 601 AGATGAAGCTTCAATTCACACCATCTGTGAGAAATGCCACCTTTCAGAGGTACCAAGGAAA 660
Qy 996 GGCAGATCCCCCGTGGCTTTGGTGGTTTCAATGCCCCCAGCATCTGTGCTTTGTGGACAG 1055
Db 661 GGCAGATCCCCCGTGGCTTTGGTGGTTTCAATGCCCCCAGCATCTGTGCTTTGTGGACAG 720
Qy 1056 CAGGTACAGCAGTGGATGGAGAGGTTTGGGCTGTACACCAGCACTTGTGCTCTGAATGA 1115
Db 721 CAGGTACAGCAGTGGATGGAGAGGTTTGGGCTGTACACCAGCACTTGTGCTCTGAATGA 780
Qy 1116 GAATGTGCTCAGTTTCAACACCTTCCAGCCACAGATTCAAAACCCAGCTCAACCTCAT 1175
Db 781 GAATGTGCTCAGTTTCAACACCTTCCAGCCACAGATTCAAAACCCAGCTCAACCTCAT 840
Qy 1176 CCACCCGG---ACATCTTCCCTCTCTCACCAGTTTCCGCTGTGAAGAGGAGGGCCCCAC 1232
Db 841 CCACCCGGGACATCTTCCCTCTCTCTCAGATTTCCTGTTAAGAAAGGAGGGCCCCC 900
Qy 1233 CCTC 1236
Db 901 ACCC 904

RESULT 8
BM990505
LOCUS AGENCOURT 8071095 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:609876
DEFINITION 5', mRNA sequence.
ACCESSION BM990505
VERSION BM990505.1 GI:22282519
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

REFERENCE
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE        NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT      Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: DCTD/DP
              cDNA Library Preparation: Rubin Laboratory
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone Distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLCM2330 row: f column: 21
              High quality sequence stop: 720.
FEATURES
source
1. .915
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /clone="IMAGE:6089876"
   /clone_lib="NIH MGC 112"
   /tissue_type="melanotic melanoma, cell line"
   /lab_host="DH10B (phage-resistant)"
   /note="Organ: skin; Vector: pOTB7; Site: 1: XhoI; Site 2:
   EcoRI; cDNA made by oligo-dr priming. Directionally cloned
   into EcoRI/XhoI sites using the following 5' adaptor:
   GGACAGAG(G). Library constructed by Ling Hong in the
   laboratory of Gerald M. Rubin (University of California,
   Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
   Superscript II RT (Life Technologies). Note: this is a
   NIH_MGC Library."
BASE COUNT      212 a 249 c 262 g 190 t 2 others
ORIGIN
Query Match      33.5%; Score 831.2; DB 14; Length 915;
Best Local Similarity 99.5%; Pred. No. 1.2e-202;
Matches 833; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1263 C T C T C T C A G T A C C A G C T C G T C C C A G G A G G A G T G G C A G A G G A T G C C A T T A T T A C T T G 1322
DB 1 C T C T C T C A G T A C C A G T C G T C C G T C C C A G A G G A G T G G C A G A G G A T G C C A T T A T T A C T T G 60
QY 1323 C A A T C T C T A G G A A T T C A T A G T T A G G C G T C G A G C T T C C A A C T T C C A C A G A G C G T G C A 1382
DB 61 C A A T C T C T A G G A A T T C A T A G T T A G G C G T C G A G C T T C C A A C T T C C A C A G A G C G T G C A 120
QY 1383 G G A G T A C A G G A G T G C C A G A G C G C C A G C C C C A G C A G A A A A A G A A G T C A G T A C C C 1442
DB 121 G G A G T A C A G G A G T G C C A G A G C G C C A G C C C C A G C A G A A A A A G A A G T C A G T A C C C 180
QY 1443 A G A A A T C A T C T T C C T T G G A A C A G G G T C T G C C A T C C C G A T G A A G A T T C G A A A T G T C A G T G C 1502
DB 181 A G A A A T C A T C T T C C T T G G A A C A G G G T C T G C C A T C C C G A T G A A G A T T C G A A A T G T C A G T G C 240
QY 1503 C A C A C T T G T C A A C A T A A G C C C C A C A C G T C T C T G T A C T G G A C T G T G G T G A G G C A C A T T 1562
DB 241 C A C A C T T G T C A A C A T A A G C C C C A C A C G T C T C T G T A C T G G A C T G T G T G T G A G G C A C A T T 300
QY 1563 T G G G C A G C T G T G C C G T C A T A C G G A C A C A G T G G A C A G G T C C T G G G C A C C C T G G C T G C 1622
DB 301 T G G G C A G C T G T G C C G T C A T A C G G A C A C A G T G G A C A G G T C C T G G G C A C C C T G G C T G C 360
QY 1623 T G T G T T T G T C C C A C T C A C G C A G A T C A C C A C A C G G C T T G C C A A G T A T C T T G C T G C A 1682
DB 361 T G T G T T T G T C C C A C T C A C G A G A T C A C A C A C G G C T T G C C A A G T A T C T T G C T G C A 420
QY 1683 G A G A A C C G C C T T G G C A T C T T T G G A A A G C G C T T C A C C C T T T G C T G T G T G T G C C C C 1742
DB 421 G A G A A C C G C C T T G G C A T C T T T G G A A A G C G C T T C A C C C T T T G C T G T G T G T G C C C C 480
QY 1743 C A A C A G C T C A A A G C C T G G C T C A G C A G T A C C A A C A A C A G T G C A G A G G T C C T G C A C C A 1802

```

```

Db 481 C A A C C A G C T C A A A G C C T G G C T C C A C A G T A C C A C A A C C A G T G C C A G G A G G T C C T G C A C C A 540
QY 1803 C A T C A G T A T A T T C C T G C C A A A T G C C T T C A G A A G G G C T G A G A T C T C A G T C C T G C A G T 1862
Db 541 C A T C A G T A T A T T C C T G C C A A A T G C C T T C A G A A G G G C T G A G A T C C C A G T C C T G C A G T 600
QY 1863 G G A A A G A T T G A T C A G T C G C T G T T G G A C A T G T A T T T G G A A G A G T T T C A C A C C T G T C T 1922
Db 601 G G A A A G A T T G A T C A G T C G C T G T T G C G A A C A T G T A T T T G G A A G A G T T T C A C A C C T G T C T 660
QY 1923 G G T G C G G C A C T C C A A G C A T G C C T T T G G C T G T G C G T G T G C A C A C C T C T G G T G G A A A G T 1982
Db 661 G G T G C G G C A C T C C A A G C A T G C C T T T G G C T G T G C G T G T G C A C A C C T C T G G T G G A A A G T 720
QY 1983 G G T C T A T T C C G G G A C A C A T G C C C T G C A G A G C T C T G T C C G A T G G G A A A G A T G C C A C 2042
Db 721 G G T C T A T T C C G G G A C A C A T G C C C T G C A G A G C T C T G T C C G A T G G G A A A G A T G C C A C 780
QY 2043 C C T C T G A T A C A T G A A G C C A C C C T G G A A G A T G T T T G G A A G A A G A A G A T G G A A A 2099
Db 781 C C T C T G A T A C A T G A A G C C A C C C T G A A G A T G T T T G G A A G A A G A A C C A G T G G G A A 837

RESULT 9
BQ057402      BQ057402      1042 bp      mRNA      linear      EST 29-MAR-2002
LOCUS         AGENCOURT_6739613 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5813014
DEFINITION    5', mRNA sequence.
ACCESSION     BQ057402
VERSION       BQ057402.1 GI:19816742
KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens
               Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 1042)
AUTHORS       NIH-MGC http://mgi.nci.nih.gov/.
TITLE         National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL       Unpublished (1999)
COMMENT       Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-r@mail.nih.gov
               Tissue Procurement: Lou Staudt
               cDNA Library Preparation: Rubin Laboratory
               cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
               Cloning Distribution: Agencourt Bioscience Corporation
               Clone Distribution: MGC clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at:
               http://image.llnl.gov row: f column: 23
               Plate: LLCM2064
               High quality sequence stop: 663.
FEATURES
source
1. .1042
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /clone="IMAGE:5813014"
   /clone_lib="NIH_MGC_99"
   /tissue_type="lymphoma, cell line"
   /lab_host="DH10B (phage-resistant)"
   /note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
   EcoRI; cDNA made by oligo-dr priming. Directionally cloned
   into EcoRI/XhoI sites using the following 5' adaptor:
   GGACAGAG(G). Size-selected >500bp for average insert size
   1.8kb. Library constructed by Ling Hong in the laboratory
   of Gerald M. Rubin (University of California, Berkeley)
   using ZAP-cDNA synthesis kit (Stratagene) and Superscript
   II RT (Life Technologies). Note: this is a NIH_MGC
   Library."
BASE COUNT      278 a 276 c 262 g 226 t
ORIGIN
Query Match      33.4%; Score 828.8; DB 14; Length 1042;
Best Local Similarity 96.9%; Pred. No. 5.4e-202;
Matches 868; Conservative 0; Mismatches 22; Indels 6; Gaps 2;

```



```
Qy 336 AATGCACTGGTCTAATGTTGGGGCTTAAGTGAATGATTCTTACTTTAAAGGAACCGG 395
Db 1 AATGCACTGGTCTAATGTTGGGGCTTAAGTGAATGATTCTTACTTTAAAGGAACCGG 60
Qy 396 GCTTCCAAAGTGTGTACTTTCTGGACCTCCACAACTGGAAAAATACCTCGAAGCAATCAA 455
Db 61 GCTTCCAAAGTGTGTACTTTCTGGACCTCCACAACTGGAAAAATACCTCGAAGCAATCAA 120
Qy 456 AATATTTTCTGGTCCATTGAAAGGAATGAACTGCTGTGGGCCCCCTCTGCCCCCAGA 515
Db 121 AATATTTTCTGGTCCATTGAAAGGAATGAACTGCTGTGGGCCCCCTCTGCCCCCAGA 180
Qy 516 ATACGAGGATGAACCATGACAGTTTACAGATCCCATACACAGTGAACAGAGAGGG 575
Db 181 ATACGAGGATGAACCATGACAGTTTACAGATCCCATACACAGTGAACAGAGAGGG 240
Qy 576 AAAGCACCACCATGGCAGAGTCCAGAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATC 635
Db 241 AAAGCACCACCATGGCAGAGTCCAGAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATC 300
Qy 636 TTCAGACTCCGAGTCCGAATGAAATGAGCCACACCTTCCACATGGTGTAGCCAGAGAAG 695
Db 301 TTCAGACTCCGAGTCCGAATGAAATGAGCCACACCTTCCACATGGTGTAGCCAGAGAAG 360
Qy 696 AGGGGTGAGGACTCTTCCCTGGTGTAGTCTTCACTGTAAGCTTCACTTAAAGAGAGG 755
Db 361 AGGGGTGAGGACTCTTCCCTGGTGTAGTCTTCACTGTAAGCTTCACTTAAAGAGAGG 420
Qy 756 AAACCTTCTTGGTGTCAAAGCAAGAGATGGGCTCCAGTGGGACAGCTGCCATCGC 815
Db 421 AAACCTTCTTGGTGTCAAAGCAAGAGATGGGCTCCAGTGGGACAGCTGCCATCGC 480
Qy 816 TCCCATCATTTGCTGTCAAGAGCGGAAAGATCACTCATGAAAGGAAGAGATTTT 875
Db 481 TCCCATCATTTGCTGTCAAGAGCGGAAAGATCACTCATGAAAGGAAGAGATTTT 540
Qy 876 GGCTGAAGAGCTGTACTCTCCAGATCTGGTGTCTGCTTTTGGTGTAGATGTC 935
Db 541 GGCTGAAGAGCTGTACTCTCCAGATCTGGTGTCTGCTTTTGGTGTAGATGTC 600
Qy 936 AGATGAAAGCTTCATTTCAACCCATCTGTGAGAAATGCCACCTTTCAAGAGTACCAAGGAAA 995
Db 601 AGATGAAAGCTTCATTTCAACCCATCTGTGAGAAATGCCACCTTTCAAGAGTACCAAGGAAA 660
Qy 996 GGCAGATCCCCCGTGGCTTGGTGTTCACATGCCCCCAGCATCTGTCTTGTGGACAG 1055
Db 661 GGCAGATCCCCCGTGGCTTGGTGTTCACATGCCCCCAGCATCTGTCTTGTGGACAG 720
Qy 1056 CAGGTACGAGCATGGATGGAGAGTTTGGGCTGACACCCAGCATTTGGTCTGGAATGA 1115
Db 721 CAGGTACGAGCATGGATGGAGAGTTTGGGCTGACACCCAGCATTTGGTCTGGAATGA 780
Qy 1116 GAACTGTGCTCAGTTCAACCTTTGCGAGCCACAAAGATTCAAACCCAGCTC--AACT 1172
Db 781 GAACTGTGCTCAGTTCAACCTTTGCGAGCCACAAAGATTCAAACCCAGCTC--AACT 840
Qy 1173 CATCCACCCGGACA--TCTTCCCCCTGCTCACCAGTTTCOGCTGTAGAAGGAGG 1225
Db 841 CATCCACCCGGACATCTTCCCCCTGCTCACCAGTTTCCCCCGTGTAAAGAAAG 896

RESULT 10
BQ064858
LOCUS BQ064858 978 bp mRNA linear EST 02-APR-2002
DEFINITION AGENCOURT 6856603 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5929312
5', mRNA sequence.
ACCESSION BQ064858
VERSION BQ064858.1 GI:19893889
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 978)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LCM2107 row: d column: 17
High quality sequence stop: 620.

FEATURES

source
1..978
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:5929312"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

BASE COUNT 258 a 261 c 244 g 213 t 2 others
ORIGIN

Query Match 33.1%; Score 821.6; DB 14; Length 978;
Best Local Similarity 98.5%; Pred. No. 3.7e-200;
Matches 849; Conservative 0; Mismatches 11; Indels 2; Gaps 2;

```
Qy 336 AATGCACTGGTCTAATGTTGGGGCTTAAGTGAATGATTCTTACTTTAAAGGAACCGG 395
Db 1 AATGCACTGGTCTAATGTTGGGGCTTAAGTGAATGATTCTTACTTTAAAGGAACCGG 60
Qy 396 GCTTCCAAAGTGTGTACTTTCTGGACCTCCACAACTGGAAAAATACCTCGAAGCAATCAA 455
Db 61 GCTTCCAAAGTGTGTACTTTCTGGACCTCCACAACTGGAAAAATACCTCGAAGCAATCAA 120
Qy 456 AATATTTTCTGGTCCATTGAAAGGAATGAACTGCTGTGGGCCCCCTCTGCCCCCAGA 515
Db 121 AATATTTTCTGGTCCATTGAAAGGAATGAACTGCTGTGGGCCCCCTCTGCCCCCAGA 180
Qy 516 ATACGAGGATGAACCATGACAGTTTACCAGATCCCCATACACAGTGAACAGAGAGGG 575
Db 181 ATACGAGGATGAACCATGACAGTTTACCAGATCCCCATACACAGTGAACAGAGAGGG 240
Qy 576 AAAGCACCACCATGGCAGAGTCCAGAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATC 635
Db 241 AAAGCACCACCATGGCAGAGTCCAGAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATC 300
Qy 636 TTCAGACTCCGAGTCCGAATGAAATGAGCCACACCTTCCACATGGTGTAGCCAGAGAAG 695
Db 301 TTCAGACTCCGAGTCCGAATGAAATGAGCCACACCTTCCACATGGTGTAGCCAGAGAAG 360
Qy 696 AGGGGTGAGGACTCTTCCCTGGTGTAGTCTTCACTGTAAGCTTCACTTAAAGAGAGG 755
Db 361 AGGGGTGAGGACTCTTCCCTGGTGTAGTCTTCACTGTAAGCTTCACTTAAAGAGAGG 420
Qy 756 AAACCTTCTTGGTGTCAAAGCAAGAGATGGGCTCCAGTGGGACAGCTGCCATCGC 815
Db 421 AAACCTTCTTGGTGTCAAAGCAAGAGATGGGCTCCAGTGGGACAGCTGCCATCGC 480
Qy 816 TCCCATCATTTGCTGTCAAGAGCGGAAAGATCACTCATGAAAGGAAGAGATTTT 875
```



```

Db 481 TCCCAATCATGCTGTCTCAGACGGGAAAGCATCACTCATGAAGGAGAGATTTT 540
Qy 876 GGCTGAAGAGCTGTGTACTCTCCAGATCCTCGTGTCTGCTTTTGTGGTGTGAATGCC 935
Db 541 GGCTGAAGAGCTGTGTACTCTCCAGATCCTCGTGTCTGCTTTTGTGGTGTGAATGCC 600
Qy 936 AGATGAAGAGCTTCAATCAACCCATCTGTGAGAAATGCCACCTTTAGAGTACCAAGGAAA 995
Db 601 AGATGAAGAGCTTCAATCAACCCATCTGTGAGAAATGCCACCTTTAGAGGTTACCAAGGAAA 660
Qy 996 GGCAGATGCCCGCTGGCTGTGGTGTTCACATGGCCCGCCAGCATCTGTCTTGTGACAG 1055
Db 661 GGCAGATGCCCGCTGGCTGTGGTGTTCACATGGCCCGCCAGCATCTGTCTTGTGACAG 720
Qy 1056 CAGGTACACAGCATGTGATGAGAGGTTTGGGCTTGACACCCAGCATCTGGTCTCTGAATGA 1115
Db 721 CAGGTACACANCANTGGATGAGAGGTTTGGGCTTGACACCCAGCATCTGGTCTCTGAATGA 780
Qy 1116 GAACTGTGCTCAGTTCACAACTCTTGGCAGCCAC-AGAATTCAACCCAGCTCAA-CTTC 1173
Db 781 GAACTGTGCTCAGTTCACAACTCTTGGCAGCCACAAAGATTTCAACCCAGCTCAAACCTTC 840
Qy 1174 ATCCACCCGACATCTTCCCCC 1195
Db 841 ATCCACCCGACATCTTCCCC 862

```

```

RESULT 11
LOCUS BO071202
DEFINITION BO071202_6853053 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:5927571
5', mRNA sequence.
ACCESSION BO071202
VERSION BO071202.1 GI:19900248
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM2102 row: 1 column: 04
High quality sequence stop: 723.
Location/Qualifiers
1. .970
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5927571"
/clone_lib="NIH_MGC_47"
/tissue_type="neuroblastoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACAGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

```

BASE COUNT 215 a 269 c 298 g 186 t 2 others

ORIGIN

```

Query Match 33.1%; Score 820.8; DB 14; Length 970;
Best Local Similarity 96.4%; Pred. No. 6e-200;
Matches 894; Conservative 0; Mismatches 23; Indels 10; Gaps 5;

Qy 1534 CTGCTACTGGAGCTGTGGTGAGGGACATTTGGGACGCTGTGCGCTCATTTACGGAGCCAG 1593
Db 1 CTGCTACTGGAGCTGTGGTGAGGGACATTTGGGACGCTGTGCGCTCATTTACGGAGCCAG 60
Qy 1594 GTGGACAGGGCTCTGGGCACCCCTGGCTGTGTTGTGTCTCCACCTGCACCCAGATCAC 1653
Db 61 GTGGACAGGGCTCTGGGCACCCCTGGCTGTGTTGTGTCTCCACCTGCACCCAGATCAC 120
Qy 1654 CACAGGGCTTGGCCAAATATCTTGTGTGAGAGAAACGGCGCTTGGCATCTTTTGGGAAAG 1713
Db 121 CACAGGGCTTGGCCAAATATCTTGTGTGAGAGAAACGGCGCTTGGCATCTTTTGGGAAAG 180
Qy 1714 CCGCTTCAACCTTTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1773
Db 181 CCGCTTCAACCTTTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 240
Qy 1774 CACAACACAGTCCAGAGGCTCTGCACACACATCAGTATGATTCCTGCCAAATGCCCTTCAG 1833
Db 241 CACAACACAGTCCAGAGGCTCTGCACACACATCAGTATGATTCCTGCCAAATGCCCTTCAG 300
Qy 1834 GAAGGGCTGAGATCTCAGTCTCAGTGGAAAGATGATCAGTTCGCTGTGGGAAACA 1893
Db 301 GAAGGGCTGAGATCTCAGTCTCAGTGGAAAGATGATCAGTTCGCTGTGGGAAACA 360
Qy 1894 TGTGATTTTGGAGAGTTTTCAGACCTGTCTGTGGCGGACATGCAAGCATGCGTTTGGCTGT 1953
Db 361 TGTGATTTTGGAGAGTTTTCAGACCTGTCTGTGGCGGACATGCAAGCATGCGTTTGGCTGT 420
Qy 1954 GCGCTGTGCACACCTCTGGCTGGAAAGTGGTCTATTCCGGGGACACCATGCCCTGCCAG 2013
Db 421 GCGCTGTGCACACCTCTGGCTGGAAAGTGGTCTATTCCGGGGACACCATGCCCTGCCAG 480
Qy 2014 GCTCTGTCCGATGGGAAAGATGCCACCTCTCTGATACATGAAGCCACCTGGAAGAT 2073
Db 481 GCTCTGTCCGATGGGAAAGATGCCACCTCTCTGATACATGAAGCCACCTGGAAGAT 540
Qy 2074 GGTTTGGAAGAGGAGTGGAAAGACACACACACACACACACACACACACACACACACAC 2133
Db 541 GGTTTGGAAGAGGAGTGGAAAGACACACACACACACACACACACACACACACACACAC 600
Qy 2134 GGGATGCGGATGAACCGCGAGTTTCAATATGCTGAACCACTTCAGCCAGCGCTATGCGAAG 2193
Db 601 GGGATGCGGATGAACCGCGAGTTTCAATATGCTGAACCACTTCAGCCAGCGCTATGCGAAG 660
Qy 2194 GTCCCCCTCTTTAGCCCCCAACTTCACGAGAGAAAGTGGAGTTTGGCTTTGACACATGAAG 2253
Db 661 GTCCCCCTCTTTAGCCCCCAACTTCACGAGAGAAAGTGGAGTTTGGCTTTGACACATGAAG 720
Qy 2254 GTCTGCTTTGGAGACTTTCCAAATATGCCCAAGCTGATTTCCCTCCCACTGAAAGCCCTGTTT 2313
Db 721 GTCTGCTTTGGAGACTTTCC-ACAAATGCCCAAGCTGATTTCCCTCCCACTGAAAGCCCTGTTT 779
Qy 2314 GCTGCGCATCTCGAGAGATGGAGGAGCGCGGAGAGAGC-GGGAGCTCGCGCAGGTGC- 2371
Db 780 GCTGCGCATCTCGAGAGATGGAGGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 839
Qy 2372 -GGGCGGCGCTCTCTGTCCA-GGGAGCTGGCAGCGCGGCC-----TGGAGGATGGGAGCGCT 2424
Db 840 GGGCGGCGCTCTCTGTCTCAGGGGAGCTGGCAGCGCGGCTTGGNAGGATGGGAGCGCTC 899
Qy 2425 CAGCAGAGAGCGGCGCCACACAGAGAG 2451
Db 900 AAGCAGAGAACGGGCGCCACACACAGAG 926

```

RESULT 12
BO059218

LOCUS BQ059218 1009 bp mRNA linear EST 29-MAR-2002
DEFINITION AGENCOURT_7049753 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5815225
5', mRNA sequence.
ACCESSION BQ059218
VERSION BQ059218.1 GI:19818558
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1009)
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ruben Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM2070 row: c column: 02
High quality sequence stop: 682.
Location/Qualifiers
1..1009
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5815225"
/clone_lib="NIH_MGC_99"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dr priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-CDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
BASE COUNT 230 a 280 c 278 g 221 t
ORIGIN
Query Match 32.3%; Score 801.8; DB 14; Length 1009;
Best Local Similarity 97.8%; Pred. No. 4.8e-195;
Matches 858; Conservative 0; Mismatches 7; Indels 12; Gaps 4;
QY 710 CTTCCCTGGTCTAGCTTTCATCTGTAGCTTCACTTAAGAGAGAACTT-CTTGGTG 768
DB 1 CTTCCCTGGTCTAGCTTTCATCTGTAGCTTCACTTAAGAGAGAACTTGGTGGTG 60
QY 769 CTCAAAGCAAGAGAGATGGGCTCCAGTTGGGACAGCTGCCATGCCATCATTTGCT 828
DB 61 CTCAAAGCAAGAGAGATGGGCTCCAGTTGGGACAGCTGCCATGCCATCATTTGCT 120
QY 829 GCTGTCAAGGACGGGAAAGACATCACTCATGAAGGAAGAG-----AGATTTTGGCT 879
DB 121 GCTGTCAAGGACGGGAAAGACATCACTCATGAAGGAAGAGATTTTGGAGATTTGGCT 180
QY 880 GAAGAGCTGTGTACTCTCCAGATCTGTGTCTGTCTTTTGTGGTGTAGATGTCAGAT 939
DB 181 GAAGAGCTGTGTACTCTCCAGATCTGTGTCTGTCTTTTGTGGTGTAGATGTCAGAT 240
QY 940 GAAAGCTTCATTCAACCCCATCTGTGAGATGCCACCTTTTCAGAGGTACCAAGGAAGGCA 999
DB 241 GAAAGCTTCATTCAACCCCATCTGTGAGATGCCACCTTTTCAGAGGTACCAAGGAAGGCA 300
QY 1000 GATGCCCCCGTGGCTTGTGGTTTCACTGGCCCCCAGCATCTGTGCTTGTGGACAGCAGG 1059
DB 301 GATGCCCCCGTGGCTTGTGGTTTCACTGGCCCCCAGCATCTGTGCTTGTGGACAGCAGG 360

QY 1060 TACCAGCAGTGGATGGAGAGGTTTGGGCTCACACCCAGCAGCTTGGTCTCTGAATGAGAAC 1119
DB 361 TACCAGCAGTGGATGGAGAGGTTTGGGCTCACACCCAGCAGCTTGGTCTCTGAATGAGAAC 420
QY 1120 TGTGCTCAGTTCAAAACCTTTCGAGCCACCAAGATTCAAAACCCAGCTCAACCTCATCCAC 1179
DB 421 TGTGCTCAGTTCAAAACCTTTCGAGCCACCAAGATTCAAAACCCAGCTCAACCTCATCCAC 480
QY 1180 CGGACATCTTCCCGCTCTCACAGATTTCGCTGTAGAAGGAGGGCCCACTCAGT 1239
DB 481 CGGACATCTTCCCGCTCTCACAGATTTCGCTGTAGAAGGAGGGCCCACTCAGT 540
QY 1240 GTGCCATGGTTTCAGGGTGAATGCTCTCAAGTACCAGCTCCGTCGCCAGAGGAGTGG 1299
DB 541 GTGCCATGGTTTCAGGGTGAATGCTCTCAAGTACCAGCTCCGTCGCCAGAGGAGTGG 600
QY 1300 CAGAGGATGTCATTATTACTTTCGCAATCTCTGAGGAATTCATAGTTGAGCGCTGCAGTT 1359
DB 601 CAGAGGATGTCATTATTACTTTCGCAATCTCTGAGGAATTCATAGTTGAGCGCTGCAGTT 660
QY 1360 CCCAATCTTCAGCAGAGCTGCAGGAGTACAGGAGGAGTGCAGGAGCGCCAGCCCA 1419
DB 661 CCCAATCTTCAGCAGAGCTGCAGGAGTACAGGAGGAGTGCAGGAGCGCCAGCCCA 720
QY 1420 GCAGAGAAAGAGTTCAGTACCCAGAAATCATCTTCTTGGAAACAGGGTCTGCCATCCG 1479
DB 721 GCAGAGAAAGAGTTCAGTACCCAGAAATCATCTTCTTGGAAACAGGGTCTGCCATCCG 780
QY 1480 ATGAAGATTCGAAATGTTCAGTGCACACTTGTCAACATAAGCCCC-GACAGCTCTCTGCT 1538
DB 781 ATGAAGATTCGAAATGTTCAGTGCACACTTGTCAACATAAGCCCCGGACACGCTCTGCT 840
QY 1539 ACTGAGCTGT-GGTGAGGGACATTTGGGAGCTGTG 1574
DB 841 ACTGAGCTGTGGTGAAGGCACATTTGGGGCCATCTG 877
RESULT 13
BQ069865
LOCUS AGENCOURT_7054570 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:5922163
DEFINITION 5', mRNA sequence.
ACCESSION BQ069865
VERSION BQ069865.1 GI:19898911
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 983)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ruben Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM2088 row: j column: 20
High quality sequence stop: 675.
Location/Qualifiers
1..983
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5922163"
/clone_lib="NIH_MGC_47"
/tissue_type="neuroblastoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:

EcORI; cDNA made by oligo-dT priming. Directionally cloned into EcORI/XhoI sites using the following 5' adaptor: GGACAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 229 a 275 c 282 g 196 t 1 others
ORIGIN

Query Match 32.2%; Score 800; DB 14; Length 983;
Best Local Similarity 95.2%; Pred. No. 1.4e-194;
Matches 869; Conservative 0; Mismatches 36; Indels 8; Gaps 4;

QY 1480 ATGAAGATTGGAATGTCTAGTCCACATTTGTAACATAAGCCCGACACGCTCTCTGCTA 1539
Db 1 ATGAAGATTGGAATGTCTAGTCCACATTTGTAACATAAGCCCGACACGCTCTCTGCTA 60

QY 1540 CTGGACTGTGTGAGGGGACATTTGGGAGCTGTGCGGTCTATTACGGAGCAGGTGGAC 1599
Db 61 CTGGACTGTGTGAGGGGACATTTGGGAGCTGTGCGGTCTATTACGGAGCAGGTGGAC 120

QY 1600 AGGGTCTGGGACACCTGGCTGCTGTGTTGTGTCCTGACCTGACGACGATCACCACAG 1659
Db 121 AGGGTCTGGGACACCTGGCTGCTGTGTTGTGTCCTGACCTGACGACGATCACCACAG 180

QY 1660 GGCCTGCCAAGTATCTTGTCTGAGAGAACGCGCTTGGCATCTTTGGAAAGCGGCTT 1719
Db 181 GGCCTGCCAAGTATCTTGTCTGAGAGAACGCGCTTGGCATCTTTGGAAAGCGGCTT 240

QY 1720 CACCCCTTGTGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1779
Db 241 CACCCCTTGTGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300

QY 1780 CAGTCCAGAGAGTCTGACACACATCAGTATGATTCCTGCCAATGCTTTCAGGAAGG 1839
Db 301 CAGTCCAGAGAGTCTGACACACATCAGTATGATTCCTGCCAATGCTTTCAGGAAGG 360

QY 1840 GCTGAGATCTCCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1899
Db 361 GCTGAGATCTCCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420

QY 1900 TTGGAAGAGTTTTCAGACTGTCTGTGCGGCACTGCAAGCATGCGTTGGCTGTGCGCTG 1959
Db 421 TTGGAAGAGTTTTCAGACTGTCTGTGCGGCACTGCAAGCATGCGTTGGCTGTGCGCTG 480

QY 1960 GTGCACACTCTGGCTGGAAGTGTCTATTCGGGGACACCATGCTCCCTGCGAGGCTCTG 2019
Db 481 GTGCACACTCTGGCTGGAAGTGTCTATTCGGGGACACCATGCTCCCTGCGAGGCTCTG 540

QY 2020 GTCCGGATGGGAAAGATGCCACCTCTCTGTATACATGAAGCCACCTGGAAGATGGTTTG 2079
Db 541 GTCCGGATGGGAAAGATGCCACCTCTCTGTATACATGAAGCCACCTGGAAGATGGTTTG 600

QY 2080 GAAAGAGAGCTGTGGAAGAACACACAGCAACAGTCCCAAGCATCAGCGTGGGGATG 2139
Db 601 GAAAGAGAGCTGTGGAAGAACACACAGCAACAGTCCCAAGCATCAGCGTGGGGATG 660

QY 2140 CGGATGAACGGGAGTTCATTATGCTGAACCACTTCAGCAGCGCTATGCCAAGTCCCC 2199
Db 661 CGGATGAACGGGAGTTCATTATGCTGAACCACTTCAGCAGCGCTATGCCAAGTCCCC 720

QY 2200 CTCCTTCAGCCCCAACTTCAGCGAG-AAAGTGGAGTTGCTTTTGACCATCATGAAGTCTG 2258
Db 721 CTCCTTCAGCCCCAACTTCAGCGAGAAAGTGGAGTTGCTTTTGACCATCATGAAGTCTG 780

QY 2259 -CTTTGGAGACTTTCACAAATGCCCAAGCTGATTCCTCCCACTG-AAAGCCCTGTTTGT 2316
Db 781 CTTTGGAGAAACTTTCACAAATGCCCAAGCTGATTCCTCCCACTGAAAGCCCTGTTTGT 840

QY 2317 GCGCATCGAGGATGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2371

Db 841 GCGGACCTTCNAAGAGAGATGAAGACCCCGAGGAGGAAAAACCGGAAGCTGCGGCAAGGG 900

QY 2372 GGGCGGCGCTCTCT 2384
Db 901 CCGGGCGGCGCTCT 913

RESULT 14
BO688069
LOCUS BO688069
DEFINITION AGENCOURT_8039759 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6210029
5' mRNA sequence.
ACCESSION BO688069
VERSION BO688069.1 GI:21813385
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 876)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2370 row: e column: 06
High quality sequence stop: 708.

FEATURES
Location/Qualifiers
1..876
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6210029"
/clone_lib="NIH_MGC_110"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pORF7; Site: 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 205 a 239 c 248 g 183 t 1 others
ORIGIN

Query Match 32.2%; Score 799.4; DB 14; Length 876;
Best Local Similarity 96.8%; Pred. No. 1.8e-194;
Matches 848; Conservative 0; Mismatches 22; Indels 6; Gaps 3;

QY 1257 TGAATGCTCTCTCAAGTACCAGCTCCGTCACGAGGAGGAGTGGCAGAGGATGCCATTAT 1316
Db 1 TGAATGCTCTCTCAAGTACCAGCTCCGTCACGAGGAGGAGTGGCAGAGGATGCCATTAT 60

QY 1317 TACTTGCAATCCTGAGGAATTCATAGTTGAGCGCTGCAGCTTCCCACTTCCAGCAGAG 1376
Db 61 TACTTGCAATCCTGAGGAATTCATAGTTGAGCGCTGCAGCTTCCCACTTCCAGCAGAG 120

QY 1377 CGTGAGGAGTACAGGAGAGTGGCAGAGCGGCCACCGCCCGCAGAGAGAAAGATCA 1436
Db 121 CGTGAGGAGTACAGGAGAGTGGCAGAGCGGCCACCGCCCGCAGAGAGAAAGATCA 180

QY 1437 GTACCCAGAAATCATCTTCCTTGGAAACAGGCTCTGCCATCCGATCGAATTCGAAATGT 1496
Db 181 GTACCCAGAAATCATCTTCCTTGGAAACAGGCTCTGCCATCCGATCGAATTCGAAATGT 240

```
Qy 1497 CAGTGCACACATTGTCAACATAAGCCCCGACACAGCTCTCTGCTACTGGAGTGTGGTGAGGG 1556
Db 241 CAGTGCACACATTGTCAACATAAGCCCCGACACAGCTCTCTGCTACTGGAGTGTGGTGAGGG 300
Qy 1557 CACATTGTGGCAGCTGTGCGGTCAATTACGAGACACAGGTGGACAGGGTCTCTGGCACCTT 1616
Db 301 CACATTGTGGCAGCTGTGCGGTCAATTACGAGACACAGGTGGACAGGGTCTCTGGCACCTT 360
Qy 1617 GGTCTGCTGTTGTGTCACCTGCACGACATCACACACGGGCTGCCAAGTATCTT 1676
Db 361 GGTCTGCTGTTGTGTCACCTGCACGACATCACACACGGGCTGCCAAGTATCTT 420
Qy 1677 GGTGCAGAGAGAACCGGCTTGGCATCTTTGGAAAGCGCTTCAACCTTTCTGCTGGTGT 1736
Db 421 GGTGCAGAGAGAACCGGCTTGGCATCTTTGGAAAGCGCTTCAACCTTTCTGCTGGTGT 480
Qy 1737 TGCCCCCAACACAGCTCAAAAGCTGTGCTCAGCAGTACCAACACAGTGCAGAGGTCCT 1796
Db 481 TGCCCCCAACACAGCTCAAAAGCTGTGCTCAGCAGTACCAACACAGTGCAGAGGTCCT 540
Qy 1797 GCACACATCAGTATGATTCCTGCCAAATGCTTCAGNAGGGGCTGAGATCTCAGTCC 1856
Db 541 GCACACATCAGTATGATTCCTGCCAAATGCTTCAGNAGGGGCTGAGATCTCAGTCC 600
Qy 1857 TGCAGTGGAAAGATTGATCAGTTCGCTGTGCGAATCATGTGATTTGGAAGAGTTCAGAC 1916
Db 601 TGCAGTGGAAAGATTGATCAGTTCGCTGTGCGAATCATGTGATTTGGAAGAGTTCAGAC 660
Qy 1917 CTGTCTGTGCGGCACTGCAAGCATCGCTTTGGCTGTGCGCTGGTGCGACACCTCTGGCTG 1976
Db 661 CTGTCTGTGCGGCACTGCAAGCATCGCTTTGGCTGTGCGCTGGTGCGACACCTCTGGCTG 720
Qy 1977 GAAAGTGTCTATTCCGGGGACACATGCCCTGCAGGCTCTGGTCC--GGATGGGGAAG 2035
Db 721 GAAAGTGTCTATTCCGGGGACACATGCCCTGCAGGCTCTGGTCCGGGATGGGGAAG 780
Qy 2036 ATGCCACCTCTCATATACATGAAGCACACCTCGAAGATGTTTGGAAAGGAGCAG-- 2092
Db 781 ATGCCACCTCTCATATACATGAAGCACACCTCGNAGATGTTTGGAAAGGAGCAG 840
Qy 2093 -TGGAAAGACACACAG-CACAACTGCCAAGCCAT 2126
Db 841 TGGGAAAGACACACAGCCACACACGTTCCAGCCAT 876

RESULT 15
BQ679122
LOCUS
DEFINITION BQ679122 928 bp mRNA linear EST 15-JUL-2002
5', mRNA sequence.
ACCESSION BQ679122
VERSION BQ679122.1 GI:21791801
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 928)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DIRP
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution by: Agencourt Bioscience Corporation
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LUCM2429 row: n column: 13
High quality sequence stop: 615.
Location/Qualifiers
```

```
source
1. .928
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6264396"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: skin; Vector: pOTB7; Site:1: XhoI; Site:2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
BASE COUNT 219 a 256 c 259 g 192 t 2 others
ORIGIN
Query Match 32.2%; Score 799; DB 14; Length 928;
Best Local Similarity 94.9%; Pred. No. 2.4e-194;
Matches 868; Conservative 0; Mismatches 42; Indels 5; Gaps 4;
Qy 1302 GAGGGATGCCATTATTAATCTTGAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCC 1361
Db 1 GAGGGATGCCATTATTAATCTTGAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCC 60
Qy 1362 CAATTCAGCAGAGCGTGCAGGAGTACAGGAGGAGTGCAGACGCGCCAGCCCCCAGC 1421
Db 61 CAATTCAGCAGAGCGTGCAGGAGTACAGGAGGAGTGCAGACGCGCCAGCCCCCAGC 120
Qy 1422 AGAGAAAGAGTCACTACCCAGAAATCATCTTCTTGGAAACAGGCTGTCATCCCCAT 1481
Db 121 AGAGAAAGAGTCACTACCCAGAAATCATCTTCTTGGAAACAGGCTGTCATCCCCAT 180
Qy 1482 GAAGATTGAAATGTCAGTGCCACATTTGTCAACATAAGCCCCGACACGTTCTGTCTACT 1541
Db 181 GAAGATTGAAATGTCAGTGCCACATTTGTCAACATAAGCCCCGACACGTTCTGTCTACT 240
Qy 1542 GGACTGTGGTGAAGGCACATTTGGGAGCTGTGCGTCATTTACGGAGACAGGTGGACAG 1601
Db 241 GGACTGTGGTGAAGGCACGTTTGGGAGCTGTGCGTCATTTACGGAGACAGGTGGACAG 300
Qy 1602 GGTCTGGGCACCTGTGCTGTGTTGTGTCACCTGCACGAGATCACACACGGG 1661
Db 301 GGTCTGGGCACCTGTGCTGTGTTGTGTCACCTGCACGAGATCACACACGGG 360
Qy 1662 CTGCGCAAGTATCTTGTGTCAGAGAGAACCGGCTTGGCATCTTTGGAAAGCGCTTCA 1721
Db 361 CTGCGCAAGTATCTTGTGTCAGAGAGAACCGGCTTGGCATCTTTGGAAAGCGCTTCA 420
Qy 1722 CCCTTTGCTGTGGTGGCCCCCAACAGCTCAAAGCTGGCTCCAGCAGTACCACCA 1781
Db 421 CCCTTTGCTGTGGTGGCCCCCAACAGCTCAAAGCTGGCTCCAGCAGTACCACCA 480
Qy 1782 GTGCCAGAGGTCTTGACACACATCAGTATGATTCCTGCCAAATGTCCTTCAGGAAGGGGC 1841
Db 481 GTGCCAGAGGTCTTGACACACATCAGTATGATTCCTGCCAAATGTCCTTCAGGAAGGGGC 540
Qy 1842 TGAGATCTCCAGTCTCTGCAGTGGAAAGATTGATCAGTTGCTGTGTTGGAAACATGTGATTT 1901
Db 541 TGAGATCTCCAGTCTCTGCAGTGGAAAGATTGATCAGTTGCTGTGTTGGAAACATGTGATTT 600
Qy 1902 GGAAGATTTCAGACCTGTCTGGTGGCGCACTGCAAGCATGCGTTGGCTGTGGCTGGT 1961
Db 601 GGAAGATTTCAGACCTGTCTGGTGGCGCACTGCAAGCATGCGTTGGCTGTGGCTGGT 660
Qy 1962 GCACACCTCT--GGCTGGAAAGTGTCTATTTCGGGGACACCATGCCCTCGCAGGCTCTGG 2020
Db 661 GCACACCTCTGGCTGGAAAGTGTGNCATTTCGGGGGACACATGCCCTCGCAGGCTCTGG 720
Qy 2021 TCCGGATGGGGAAGATGCCACCTCTCTGATACATGAAGCAACCTCGAAGATGTTGG 2080
Db 721 GCCGGATGGGGAAGATGCCACCTCTCTGATACATGAAGCAACCTCGAAGATGTTGG 780
```


